THIS BOOK IS A PART
OF THE LIBRARY OF



THE TRVE VNIVERSITY IS A
COLLECTION OF BOOKS. CARLYLE

Dr. A. K. SAHA.
NUCLEAR PHYSICS DIVISION.
SAHA INSTITUTE OF NUCLEAR PHYSICS,
92, Acharya Prafulla Chandra Road,
CALCUTTA-9.

DR. AJIT KUMAR SAH
125, SOUTHERN AVENUE,
CALCUTTA-29

A. K. SAHA.
NUCLEAR PHYSICS DIVISION.
SAHA INSTITUTE OF NUCLEAR PHYSICS,
92, Acharya Prafulla Chandra Road,
CALCUTTA-9.

ELECTRON OPTICS AND THE ELECTRON MICROSCOPE

BOOKS BY V. K. ZWORYKIN
AND COLLABORATORS

ELECTRON OPTICS AND THE
ELECTRON MICROSCOPE

By V. K. Zworykin, G. A. Morton, E. G. Ramberg, J. Hillier, A. W. Vance; *all at the RCA Laboratories, Princeton, New Jersey. 766 pages, 5⅝ by 8⅝. 578 illustrations. Cloth.*

TELEVISION: THE ELECTRONICS OF IMAGE TRANSMISSION

By V. K. Zworykin and G. A. Morton. *646 pages, 6 by 9⅛. 509 illustrations. Cloth.*

PHOTOCELLS AND THEIR APPLICATIONS

By V. K. Zworykin and E. D. Wilson, *Westinghouse Electric and Manufacturing Company. Second Edition. 348 pages, 5⅝ by 8⅝. 180 illustrations. Cloth.*

PUBLISHED BY
JOHN WILEY & SONS, INC.

ELECTRON OPTICS AND THE ELECTRON MICROSCOPE

BY

V. K. ZWORYKIN, E.E., Ph.D.
RCA Laboratories
Princeton, N. J.

G. A. MORTON, Ph.D.
RCA Laboratories
Princeton, N. J.

E. G. RAMBERG, Ph.D.
RCA Laboratories
Princeton, N. J.

J. HILLIER, Ph.D.
RCA Laboratories
Princeton, N. J.

A. W. VANCE, E.E.
RCA Laboratories
Princeton, N. J.

NEW YORK
JOHN WILEY & SONS, INC.
LONDON: CHAPMAN & HALL, LIMITED



THIS BOOK HAS BEEN MANUFACTURED IN ACCORDANCE WITH THE RECOMMENDATIONS OF THE WAR PRODUCTION BOARD IN THE INTEREST OF THE CONSERVATION OF PAPER AND OTHER IMPORTANT WAR MATERIALS.

COPYRIGHT, 1945
BY
VLADIMIR K. ZWORYKIN, GEORGE A. MORTON, EDWARD G. RAMBERG, JAMES HILLIER, AND ARTHUR W. VANCE

All Rights Reserved
This book or any part thereof must not be reproduced in any form without the written permission of the publisher.

PRINTED IN THE UNITED STATES OF AMERICA

A. K. SAHA.
PHYSICS DIVISION.
OF NUCLEAR PHYSICS,
Acharya Jagadis Chandra Bose Road,
CALCUTTA-9.

PREFACE

Electron microscopy, which, within the past ten years, has developed from a subject of purely academic interest into one of great practical importance, is engaging the attention of men in all branches of science. Recognizing this, the authors have for some time felt that there was a real need for a comprehensive text covering the electron microscope in all its phases.

The material for the book was chosen to fulfil a twofold purpose. The first is to aid the present or prospective electron microscopist in understanding his instrument and in using it to greatest advantage; the second, to present systematically the practical and theoretical knowledge which must form the basis for further progress in electron microscope design. To this end the book has been divided into two parts. The first part contains descriptions of various types of electron microscopes together with a non-mathematical discussion of the electron optical theories on which the electron microscope is based and the practical information necessary for its effective operation. The second part presents a survey of theoretical electron optics and employs mathematics as liberally as a methodical development of the subject matter warrants. This treatment is intended to supplement the practical information of the first part and may serve as a guide in the electron optical design of improved instruments.

The first part of the book opens with a qualitative introduction to the principles of electron optics and a survey of its applications. These sections are succeeded by a description of the different types of electron microscopes, stressing instruments other than the magnetic electron microscope. Next the design principles of the several components of the magnetic electron microscope are outlined, the most successful instruments are described in detail, and the operation tolerances of the magnetic electron microscope are given. A chapter is devoted to the construction of suitable electric power supplies for the microscope. The first part closes with a chapter on the techniques of electron microscopy and a general survey of the research accomplishments of the new instrument.

The second part of the book begins with a discussion of the theoretical basis of electron optics. Chapters dealing with the measurement and calculation of electrostatic fields, the tracing of electron rays through

such fields, and the properties of various types of electrostatic lenses (and mirrors) follow in sequence. They are succeeded by analogous treatments of magnetic fields and magnetic lenses. The aberrations of electron lenses are derived in systematic fashion and discussed quantitatively, possible ways of correcting them being pointed out. The modifications introduced by the variation of the electron mass with its velocity at high accelerating potentials, as well as those due to the greater mass of ions in ion optics, are treated in a later chapter. The final chapter attempts to summarize our present knowledge of the process of image formation in the electron microscope. A brief discussion of noise problems arising in connection with electron multipliers and with the scanning microscope, as well as a few useful tables, form an appendix to the book.

Throughout the preparation of this book the authors have benefited from the assistance and helpful criticism of their associates in the RCA organization, to whom they gratefully acknowledge their indebtedness.

V. K. Zworykin
G. A. Morton
E. G. Ramberg
J. Hillier
A. W. Vance

Princeton, New Jersey
August 1945

CONTENTS

PART I. PRACTICAL ELECTRON OPTICS AND ELECTRON MICROSCOPY

CHAPTER

1. ELECTRON OPTICS
   1. Electrons and Electron Emission 1
   2. Analogy of Electrons and Light 4
   3. Motion in Electric Fields; Electric Electron Lenses 8
   4. Motion in Magnetic Fields; Magnetic Electron Lenses . . . 14
   5. General Properties of Electron Lenses; Electron Mirrors . . . 18
   6. Image Defects of Electron Lenses 21
   7. High-Voltage Electron Optics; Ion Optics 30

2. APPLICATIONS OF ELECTRON OPTICS
   1. Electron Guns 33
   2. Image-Reproducing Devices 42
   3. Deflection Systems 52
   4. Beam Tubes and Multipliers 55
   5. Velocity and Mass Spectrographs 64
   6. Particle-Accelerating Devices 71

3. ELECTRON MICROSCOPES
   1. Reason for Electron Microscope 81
   2. Magnetic Microscope: Light-Optical Analogue 86
   3. The Electrostatic Microscope 89
   4. Scanning Microscope 98
   5 Point and Shadow Microscopes 106
   6. X-ray Shadow Microscope 111
   7. Emission Microscopes 112

4. ELECTRON OPTICS OF HIGH MAGNIFICATION
   1. Factors Influencing Resolving Power 123
   2. Electron Sources 125
   3. The Condenser Lens 131
   4. The Objective 134
   5. The Projector 147
   6. Depth of Field 155
   7. Object Thickness and Resolution 156
   8. Object Thickness and Contrast 158
   9. Electrical and Magnetic Disturbances 159

5. THE MAGNETIC ELECTRON MICROSCOPE
   1. General Construction 163
   2. The Electron Source 177
   3. The Condenser, Objective, and Projector Lenses 181
   4. Object Chamber 185
   5. Plate Chamber 190

CHAPTER

6. The Vacuum System 195
7. Controls and Adjustments 199
8. Modifications for High-Voltage Operation 200

6. ABERRATIONS AND TOLERANCES IN THE ELECTRON MICROSCOPE

1. Spherical Aberration and Diffraction 205
2. Chromatic Aberration 206
3. Field Aberrations 208
4. Current and Voltage Stabilization Tolerances 213
5. Construction and Alignment Tolerances 216
6. Shielding Requirements 217
7. X-ray Protection 219

7. ELECTRON MICROSCOPE POWER SUPPLIES

1. General Requirements 223
2. High-Potential Supply 223
3. Radio-Frequency Circuits 228
4. Rectifier Circuits 231
5. Regulating Circuits 233
6. Magnetic-Lens Current Supplies 236
7. Filament Supplies 238

8. MANIPULATION OF THE ELECTRON MICROSCOPE

1. Preparation of the Specimens 241
2. Alignment of Magnetic Electron Microscope 252
3. Focusing 255
4. Determination of the Magnification 258
5. Sensitive Materials 259
6. Electron Stereomicroscopy 262
7. Surface Study with Reflected Electrons 269
8. Electron Diffraction Patterns 270

9. THE ELECTRON MICROSCOPE AS A RESEARCH INSTRUMENT

1. Present Range of Application of the Electron Microscope . . 281
2. Biological Research 282
3. Chemical Research 310
4. Metallurgy 331
5. Prospects 337

PART II. THEORETICAL BASIS OF ELECTRON OPTICS AND THE ELECTRON MICROSCOPE

10. THEORETICAL BASIS OF ELECTRON OPTICS

1. Basic Problems of Electron-Optical Design 343
2. The Electron — Particle and Wave Aspects 344
3. Electron Optics of Electrostatic Fields 351
4. Electron Optics of Magnetic Fields 353
5. Intensity Relations in Electron Beams and Electron Images . 354

CHAPTER

11. DETERMINATION OF POTENTIAL DISTRIBUTION

1. The Laplace Equation 358
2. Two-Dimensional Systems 359
3. Potential Distribution by Means of Complex Functions . . . 363
4. Problems Involving Separation of the Variables 368
5. Special Relationships in the Two-Dimensional Field 371
6. Axially Symmetrical Systems 375
7. Typical Configurations with Axial Symmetry 378
8. Selection of Coordinates 383
9. Graphical Potential Mapping: The Liebmann Procedure . . . 386
10. Potential Mapping by Means of the Electrolytic Plotting Tank 389
11. Three-Dimensional Potential Distributions 396
12. Potential Distribution in the Presence of Space Charge . . . 396

12. ELECTRON TRAJECTORY TRACING

1. The Ray Equation 400
2. Analytical Solution of the Ray Equation 404
3. Numerical Methods; Differences 407
4. Path Integration in Approximated Fields (Gans) 412
5. Graphical Methods 414
6. Mechanical Path-Plotting Mechanisms 417
7. The Rubber Model 418

13. GAUSSIAN DIOPTRICS OF ELECTROSTATIC LENSES

1. Axial Symmetry and Lens Action 423
2. Cardinal Points 426
3. Fundamental Theorems Regarding Electron Lenses 429
4. Classification of Electron Lenses 433
5. The Thin Lens 435
6. Unipotential Lenses 437
7. Single-Aperture Lenses 441
8. Immersion Lenses 447
9. Cathode Lenses 452
10. Electron Mirrors 456

14. MAGNETIC FIELDS

1. Determination of Fields in the Absence of Iron 466
2. Magnetic Fields in the Presence of Iron 475
3. Magnetic Shielding 481
4. Permanent Magnets 484
5. Measurement of Field Distributions 486

15. ELECTRON MOTION IN MAGNETIC FIELDS AND MAGNETIC LENSES

1. General Properties 493
2. Motion in Uniform Magnetic Field 495
3. Motion in Axially Symmetric Magnetic Fields 499
4. Approximate Methods of Path Determination in Axially Symmetric Magnetic Fields; the Thin Lens; Magnetic Lenses without Image Rotation 512

CHAPTER

5. Electron Motion in "Two-Dimensional" Magnetic Fields . . 515
6. Superposed Electric and Magnetic Fields 520
7. Combined Electric and Magnetic Lenses 535

16. ABERRATIONS OF ELECTRON LENSES

1. Classification of Aberrations 540
2. The Derivation of the Geometrical Aberrations 541
3. Evaluation of the Geometrical Aberration Constants 555
4. Validity of the Aberration Expressions 563
5. Geometric Aberrations of the Electron Mirror 564
6. Chromatic Aberrations of Electron Lenses 570
7. Chromatic Aberrations of the Electron Mirror 575
8. Aberrations of Cathode Lenses 577
9. Space-Charge Effects 591
10. Instabilities, Disturbances, and Misalignments 595

17. MAGNITUDE AND CORRECTION OF ELECTRON LENS DEFECTS

1. Approaches to the Aberrations Problem 603
2. The Aperture Defect 603
3. Coma 633
4. Curvature of Field and Astigmatism 635
5. Distortion 639
6. Chromatic Aberration 641
7. Space-Charge Defect 647

18. HIGH-VOLTAGE ELECTRON OPTICS. ION OPTICS

1. Magnetic Lenses 650
2. Electrostatic Lenses. 651
3. Compound Lenses at High Voltages; Aberrations 656
4. Ion Optics 658
5. Relativistic and Focusing Effects in Accelerating Devices . . . 660

19. IMAGE FORMATION IN THE ELECTRON MICROSCOPE

1. The Mechanism of Image Formation 672
2. Scattering and Absorption Processes in the Object 674
3. Image Formation with Small Aperture 691
4. Image Formation with Large Aperture 701
5. Contour Effects 708
6. Resolution and Object Thickness 720
7. The Limit of Resolution of the Electron Microscope 724
8. Limits of the Recognition of Small Objects; Dark-Field Operation 729
9. Crystalline Diffraction Effects 733
10. Limits of Resolution of Other Types of Electron Microscopes . 735

APPENDIX

I. THE PROBLEM OF NOISE IN AMPLIFICATION AND IN THE SCANNING MICROSCOPE

1. Reduction of Noise by the Employment of an Electron Multiplier as Preamplifier 747
2. Noise in the Scanning Microscope 749

APPENDIX

II. MISCELLANEOUS TABLES

1. Fundamental Physical Constants 753
2. Some Useful Derived Constants 753
3. Conversion Factors 753
4. Mean Free Path of Electrons in Air as Function of Pressure . . 754
5. De Broglie Wave Lengths of Various Particles in Motion . . . 754

AUTHOR INDEX 755

SUBJECT INDEX 759

NOTE

The metric centimeter-gram-second system of units has been adopted throughout. Electrical quantities (charge, current, voltage, resistance) are expressed in electrostatic units (e.s.u.), magnetic quantities (magnetic potential and magnetic field) in electromagnetic units (e.m.u.). Practical units are employed for these quantities in all equations in which constant factors have been reduced to numerical coefficients.

SAHA.
CS DIVISION.
CLEAR PHYSICS,

PART I. PRACTICAL ELECTRON OPTICS AND ELECTRON MICROSCOPY

CHAPTER 1

ELECTRON OPTICS

1·1. Electrons and Electron Emission. The electron is one of the basic constituents of matter. For many purposes it may be thought of as a small electrically charged sphere with a diameter of the order of 10^{-12} cm, a charge of $1.60 \cdot 10^{-19}$ coulomb, and a mass of $9 \cdot 10^{-28}$ gram. Every atomic nucleus is surrounded by an atmosphere of these minute charged particles coursing about it in planetary orbits, the individual electrons being prevented from escaping by the strong attracting electric

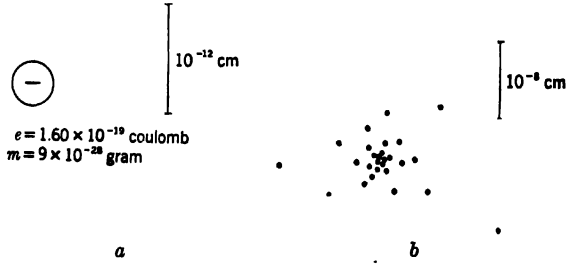


FIG. 1·1. Relative Dimensions of an Electron and an Atom. (*a*) The Electron. (*b*) Electron Atmosphere of an Atom (Copper).

field of the positively charged nucleus. This attractive force is least for the electrons traveling in the outer fringe of the atmosphere, being here in large part balanced by the repulsive forces exerted by the remaining electrons. These outer, or valence, electrons are hence more readily detached from the atom than the rest. An idea of the relative dimensions of an electron and an atom, as well as of the distribution of electrons in the atmosphere surrounding an atomic nucleus, is given by Fig. 1·1.

When a large number of atoms of the so-called metallic elements —

for which the binding of the outermost electrons is especially loose — aggregate and form solid substance, that is, a block of metal, the attracting force exerted on these valence electrons by neighboring atomic cores or *ions* becomes comparable to that of the parent core, so that these valence electrons pass readily from the domain of one to the next. They become so-called *free electrons* within the metal, free to move through the entire block (Fig. 1·2). If the block is subjected to an electric field, they will drift in the direction of the force exerted on them

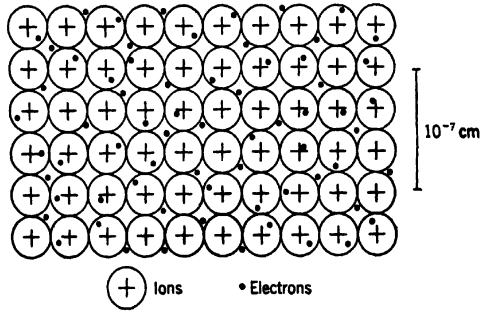


FIG. 1·2. Positive Ions and Free Electrons Composing a Metal (for Example, Copper).

by the field and thus give rise to an electric current in the metal, while the heavier atomic ions remain fixed in the regular crystalline framework of the metal. Collisions between these moving electrons and the ions will cause part of the kinetic energy of the electrons to be converted into heat, that is, energy of vibration of the ions. This gives rise to the *resistance heating* of a wire carrying current.

Under normal circumstances the free electrons cannot leave the metal; the attractive force of the positively charged atom cores or ions near the surface prevents their escape. Normally their velocity within the metal is too small to overcome this attraction. There are, however, a number of ways of getting electrons out of metals into free space. Perhaps the most important of these consists in heating the metal to a high temperature, that is, setting the ions of the metal into strong vibrations. Collisions between free electrons and ions, as well as of the former among themselves, give some of the electrons sufficient energy to escape through the surface.

The heating itself may conveniently be accomplished by passing a current through the metal, as shown in Fig. 1·3. Here the electron-emitting metal has been shaped into a thin filament so as to achieve a

high current concentration and consequently a high temperature with a moderate current. If, now, this filament were isolated within a glass bulb, the escape of electrons would soon cease, since the filament would

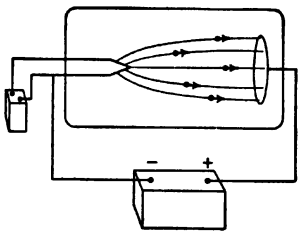


FIG. 1·3. Thermionic Emission of Electrons.

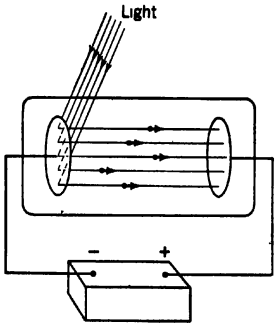


FIG. 1·4. Photoemission of Electrons.

become more and more charged up positively, exerting a greater attractive force on the escaping electrons and causing their return to the filament. Hence it is necessary to place in the bulb another piece of

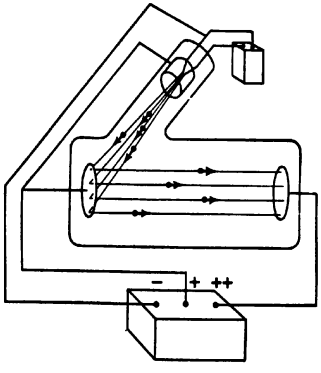


FIG. 1·5. Secondary Emission of Electrons.

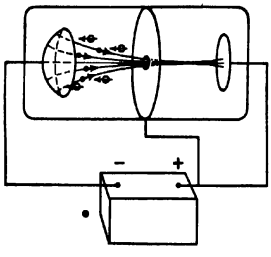


FIG. 1·6. Emission of Electrons in Gas Discharge.

metal, connected to the filament (for example, through a battery as shown) so as to create an electric field which draws off the electrons from the filament to the plate or *anode*.

Electrons within the metal may also be given sufficient velocity to

leave the metal by illuminating the surface of the metal with light (Fig. 1·4). The energy of the light absorbed by the metal, being concentrated in individual energy packets, light quanta or *photons*, is transferred to some of the metallic electrons, which thus can leave the metal and eventually be collected by an anode. Quite similarly (Fig. 1·5), a stream of fast (primary) electrons impinging on a metal plate can give the metal electrons sufficient energy to escape, giving rise to the emission of *secondary electrons*. The electrons may also obtain the required energy from positive ions, or positively charged gas atoms or molecules, striking the metal as shown in Fig. 1·6. For this purpose a slight amount of gas is left in the region between the anode and the electron emitter or *cathode*; the former may be provided with a small hole to permit the escape of the electrons into a more highly evacuated region. The positive ions are initially produced by the removal of the most loosely bound electrons from gas atoms or molecules in impacts with rapidly moving electrons or other ions, the latter being accelerated by a strong electric field between the cathode and the anode. A sufficient number of charged particles is always present in a gas to start off the *discharge*.

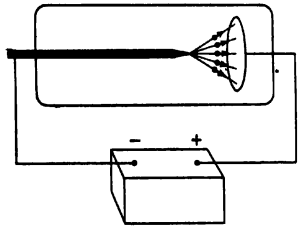


FIG. 1·7. Field Emission of Electrons.

One more method for releasing electrons from metals relies on reducing the effectiveness of the attractive force of the metal ions near the surface by the application of a strong external opposing field (Fig. 1·7). The opposing field must, for this purpose, be of the order of several million volts per centimeter and can best be attained by shaping the cathode into an extremely sharp point. The electron emission so obtained is usually referred to as *field emission* or *cold emission*.

1·2. Analogy of Electrons and Light. Electron optics deals with the propagation of electrons as light optics deals with that of light. Just as a light beam may be conceived of as a wave motion guiding minute packets of energy or mass, that is, light quanta or photons, an electron beam may be represented as a wave guiding the individual electrons. As will be seen in greater detail later, the laws governing light optics and electron optics show a close resemblance. Apart from this, deep-seated, qualitative differences characterize the behavior of electrons and of light.

The distinction which has made electrons particularly useful in *microscopy*, the examination of very small objects, is quantitative rather

than qualitative: The wave length of an electron wave (for electrons traveling at the velocity attained when accelerated through a difference of potential of about 50 kilovolts) is only about 1/100,000 as long as that of a wave of visible light. The significance of this is appreciated when light waves and sound waves are compared, the latter having wave lengths of the order of a million times as long as the former.

In the absence of external influences sound waves, light waves, and electron waves are all propagated in straight lines. However, if an obstacle is placed in their path, they will, to some extent, bend around the obstacle, into the latter's shadow. The degree of bending depends upon the wave length, decreasing as the latter is decreased. Thus, a man standing behind the wall at B in Fig. 1·8 can hear sound from the

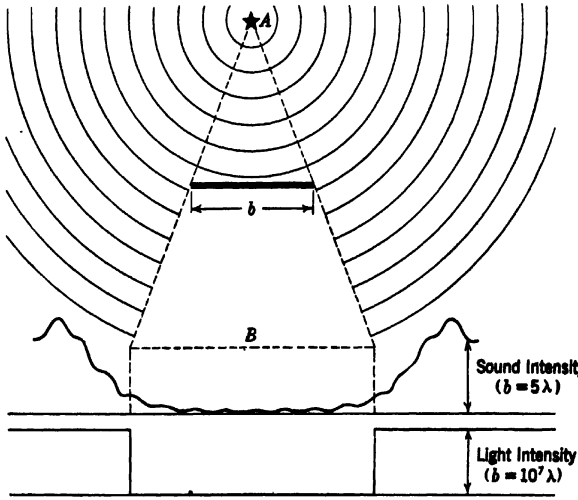


FIG. 1·8. Diffraction of Sound and Light Waves by a 20-Foot Wall.

source placed at A, but cannot see light from a similarly placed light source. This is a familiar phenomenon: Sound does not cast sharp shadows, as does light. In fact, to achieve a bending of light waves about an obstacle similar to that of the sound waves, all dimensions would have to be reduced by a factor of a million. With electrons, having a much smaller wave length, the bending, or *diffraction*, of the waves plays a still smaller role. To imitate the condition obtaining for light the obstacle would have to be reduced by a factor of a hundred thousand! The propagation of electrons, and hence the degree of dif-

fraction, is, for velocities small compared to that of light, inversely proportional to their velocity or the square root of the accelerating voltage.

Now, in the design of ordinary light-optical instruments, such as telescopes and microscopes, light is not treated as a wave motion, but as consisting of a set of mutually independent light *rays*, which in the wave picture correspond to the normals of the wave fronts (Fig. 1·9).

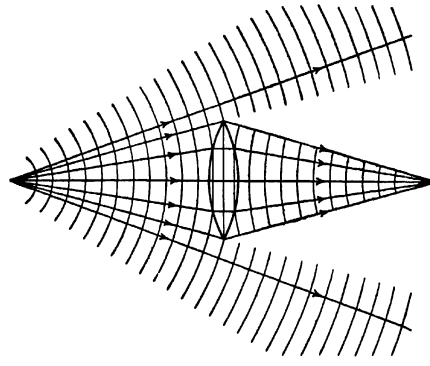


FIG. 1·9. Refraction of Light Waves and Light Rays by a Lens.

The paths of these rays depend only on their point of origin and initial direction and on the variation of the index of refraction along their path. In regions of uniform index they are straight lines, whereas at boundaries between different regions they are reflected, the reflected ray making the same angle with the normal to the boundary surface as the incident ray, or refracted, the ratio of the sines of the angles of incidence and refraction being given by the inverse ratio of the corresponding indices of refraction in the two media, as is required by Snell's law, or both (Fig. 1·10). W. R. Hamilton, over a century ago, noted that the course of these light rays was controlled by the same laws as the paths of material particles acted on by conservative forces (derivable from a potential). Examples of such forces are those exerted by a gravitational field,

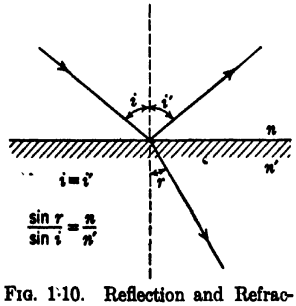


FIG. 1·10. Reflection and Refraction of a Light Ray.

or by an electric field if the particle is charged. The possible paths of the light rays and the particles will be identical if, all through space, the index of refraction in the former case is proportional to the velocity of the particle in the latter. It is generally convenient to represent the velocity of the electron by the square root of the potential ϕ, the latter being measured from a point where the electron velocity vanishes. This emphasizes the fact that, for any given electron, this velocity is, like the index of refraction for a particular light ray, a function of position only. There is thus an exact correspondence between the *ray optics* or *geometrical optics* of the light-optical designer and the Newtonian particle mechanics. It is to be expected, however, that, in view of the much shorter electron wave length, *electron-optical* instruments designed on the basis of ordinary particle mechanics will function in accord with their design even when dealing with objects of such small dimensions that light-optical instruments have become inapplicable.

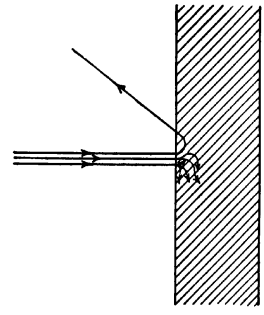


FIG. 1·11. Stopping an Electron Beam by Matter.

Among the qualitative differences, the most important deviation in the behavior of electrons from that of light consists in the fact that electrons can be deflected by electric and magnetic fields, whereas light, under normal circumstances, is not influenced by either.

The circumstance that matter is quite generally opaque to electrons may be regarded as a consequence of their being deflected by an electric field. The strong local fields of the positively and negatively charged elementary particles making up matter deflect electrons passing at close range and may cause them, eventually, to give up part of their kinetic energy (Fig. 1·11). Some of the original electrons may ultimately be scattered backwards out of the surface with a velocity comparable with the original velocity. Most of them, however, gradually lose their velocity in collisions with the constituents of the material and come to form part of the negative charge of the substance. Although relatively thick layers of matter are needed to stop a high-speed electron beam completely (for 50-kilovolt electrons 0.1 millimeter of aluminum or about 20 centimeters of air are required for this purpose), even one ten-thousandth of a millimeter of aluminum or a fifth of a millimeter of atmospheric air will cause an originally parallel electron beam to spread out into one with a mean aperture angle of 10 degrees (Fig. 1·12). Thus all vessels in which electrons are to travel must be carefully

evacuated, and the use of material media, such as glass, for electron lenses is out of the question.

However, as will be seen in greater detail below, electron lenses can be realized because of the possibility of deflecting electrons by electric and magnetic fields, a phenomenon which was utilized as early as 1897 by J. J. Thomson in his work leading to the discovery of the electron. Thus the very effect which has been seen to be responsible for the inapplicability of material lenses provides a substitute for them.

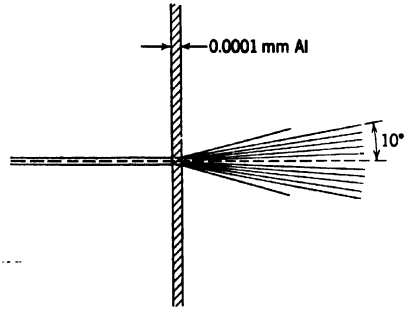


FIG. 1·12. Spreading of an Electron Beam After Passage through a Thin Film.

1·3. Motion in Electric Fields; Electric Electron Lenses. The possibility of utilizing the deflection of electrons by electric fields for the design of electron-optical instruments is implicit in the earlier statement that electron paths in electric fields are identical with light paths in a refractive medium, provided that the refractive index at every point be made proportional to the square root of the potential (and hence to the velocity of the particle) at the corresponding point in the electric field.

This may now be demonstrated with the aid of some examples. To begin with, consider refraction and total reflection at the interface between two regions at different potential. Let the region of higher potential ϕ_2 be separated from that of lower potential ϕ_1 (Fig. 1·13) by a pair of slightly separated, parallel metal plates A and B. Each plate is assumed to be provided

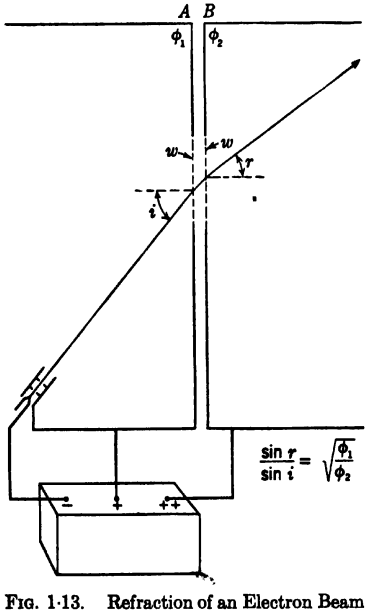


FIG. 1·13. Refraction of an Electron Beam Passing into a Region at Higher Potential.

with a *Lenard window*, a window of aluminum foil thin enough to permit electrons to traverse it with a minimum of absorption and scattering. In the field-free regions to the left of plate A and to the right of plate B the electrons will travel along straight lines. After passing through the window in plate A, however, they encounter a strong accelerating field, bending their paths, so that they meet the window in plate B more nearly in a normal direction. They retain this direction as they pass through the window out into the space to the right of B. The behavior of the electrons is here entirely analogous to the behavior of light rays passing from an optically less dense to an optically more dense medium. Only in one respect is there a material difference; there is no simultaneously reflected electron ray. This is due to the fact that in the electric model the change in potential, or index of refraction, is gradual. If it is made similarly gradual in the optical case, no reflected ray will result.

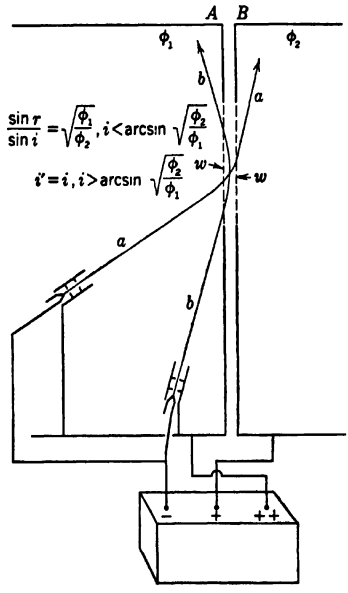


FIG. 1·14. Refraction and Reflection of an Electron Beam at Interface between Regions of Higher and Lower Potential.

In the case of a transition from a higher-potential to a lower-potential region (Fig. 1·14) the field between the plates retards the electrons. If the angle of incidence of the latter is sufficiently large they are turned back before reaching plate B. The condition for this happening is identical with that for the occurrence of total reflection of light meeting an interface between an optically denser or less dense medium. For smaller angles of incidence, for which refraction takes place (that is, the incident ray reaches B), the reflected ray is again absent.

Even in the present, rather artificial, example the sharp boundary between media of different refractive index so characteristic of light-optical devices has been replaced by a region of gradually changing index, that is, by a uniform electric field. That is typical of electron optics; in fact, even if for practical reasons only, the use of material conductors, such as the Lenard windows, in the path of the electrons is normally

avoided entirely. Since these alone can bring about abrupt changes in the variation in space of the electron-optical index, electron optics deals only with media of continuously variable optical density.

Similar light-optical media are found in nature. Thus the atmosphere of the earth decreases continuously in density and refractive index from the ground up, giving rise to the phenomenon of astronomical refraction. As shown in Fig. 1·15, the resulting bending of the rays from

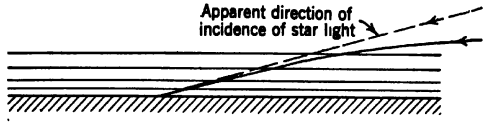


FIG. 1·15. Astronomical Refraction in the Atmosphere (Schematic).

a star near the horizon causes such a star to appear higher in the sky than it is in fact. The fine lines indicate surfaces of equal density and index of refraction. The case of astronomical refraction is seen to be closely analogous to the deflection of an electron beam by an ordinary electric deflecting field (Fig. 1·16). Here again the fine lines indicate

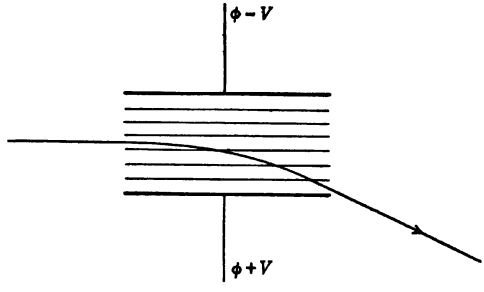


FIG. 1·16. Action of Uniform Electric Deflecting Field.

surfaces of equal index or electrical potential, in short, equipotential surfaces. The actual paths of the electrons can be determined approximately by applying the law of refraction at each equipotential surface, regarding the intervening regions as being at uniform potential.

Again, the crystalline lens of the human eye (Fig. 1·17) consists of a multiplicity of layers whose index increases toward the center. A typical electron lens, such as that formed by the field between coaxial cylinders at different potentials (Fig. 1·18), shows a close similarity to it. This resemblance is not complete, however, since surfaces of con-

stant index, or equipotentials, cannot be closed surfaces in the case of an electron lens. The added difference that, in the electron lens shown, the index of refraction, or potential, increases uniformly in one direction does not apply generally.

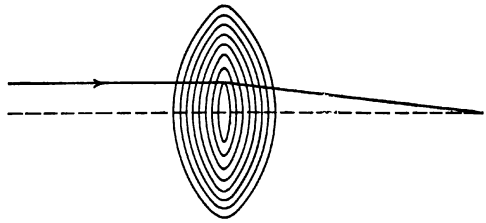


FIG. 1·17. The Crystalline Lens of the Human Eye.

Similar electric electron lenses are formed by any set of electrodes shaped so as to be symmetrical about a common axis. They all share the following characteristics: In any plane, for a region close to the axis, the electric field exerts on an electron a force component toward or away from the axis which is proportional to the distance of the electron from the axis (Fig. 1·19). To the same approximation the force acting in

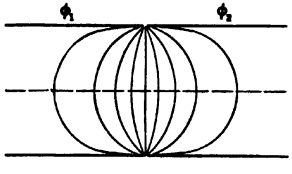


FIG. 1·18. Electric Electron Lens Formed by Two Coaxial Cylinders at Different Potentials.

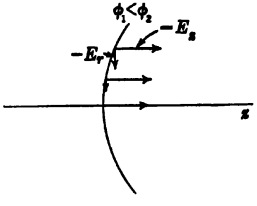


FIG. 1·19. Variation of Forces Acting on an Electron with Distance from Axis.

the direction of the axis may be regarded as the same throughout the region considered on any equipotential. As a result the total change in slope, or the total deflection, of rays incident from a common center of divergence (object point) on the axis at different heights on the lens is proportional to the height of incidence (and emergence); the rays are hence reunited in a new point on the axis, the image point, or diverge in such a manner that they appear to come from a point on the axis (the *virtual* image point) (Fig. 1·20).

Optically the action of the elementary regions between equipotentials — like that of an ordinary glass lens — may be compared to that of a

prism, whose effective vertex angle, and hence ray-deflecting power, increases linearly with the distance from the axis (Fig. 1·21). The lens is obtained by rotating a thin element of such a prism about the optic axis.

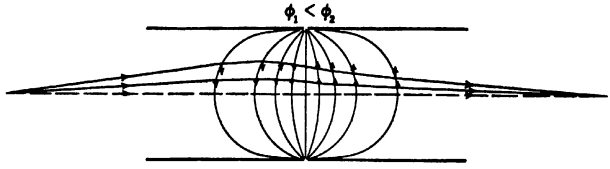


FIG. 1·20. Focusing Action of an Electric Lens on Electrons Diverging from a Point on the Axis.

A lens of the type considered images not merely one point on the axis into another, but does the same for points adjoining the *object point* and lying in the same plane normal to the axis. The images of the latter are located relative to the image point on the axis in the same manner as the off-axis object points relative to the axial object point. Thus a small object in the object plane and near the axis is faithfully imaged in the *image plane* (Fig. 1·22). The magnification of the image varies with the position of the object plane, increasing as the object is moved up toward the lens and the image moves away from the lens correspondingly. If the image is real it will, in general, be found to be inverted as shown in the figure.

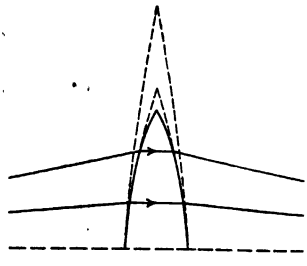


FIG. 1·21. Lens Element as Prism of a Variable Effective Vertex Angle.

For a so-called *thin lens*, that is, a lens whose refracting field is very short compared to the distance of the lens both from the object and

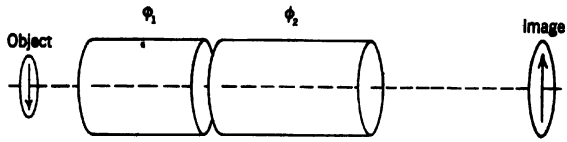


FIG. 1·22. Image Formation by an Electric Electron Lens.

from the image, the location and size of the image can be constructed very simply if only the two focal points F_i and F_o are given. The image-side focal point F_i is the point on the axis to which the rays

parallel to the axis on the object side converge; the object-side focal point F_o, that at which rays parallel to the axis on the image side converge (Fig. 1·23). As the paths of the rays may be traveled in either direction, the object-side focal point may also be defined as the point of origin of rays emerging from the lens parallel to the axis.

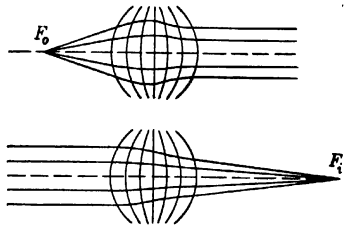


FIG. 1·23. Definition of the Focal Points F_o and F_i.

The location of the image of any off-axis object point P is found simply by drawing the ray through P and F_o, which must emerge parallel to the axis, and that through P incident parallel to the axis, which must pass through F_i (Fig. 1·24). The intersection of the two rays locates the image point P' and thus establishes both the location and the magnification of the image.

In only a few cases can an electron lens be considered thin in the above sense. In the more usual instance of a thick lens — or a refractive field of considerable extent — it is necessary to introduce two more elements, the object-side and image-side principal planes H_o and H_i, to define

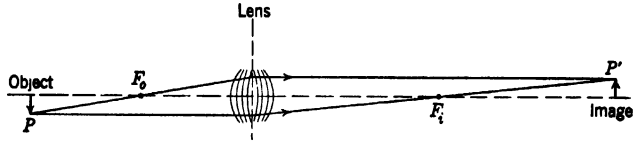


FIG. 1·24. Determination of Image Position and Magnification for a Thin Lens.

fully the lens action of the field. The principal planes are the two virtual planes which the lens images into each other with unity magnification. In other words, a ray directed from object space at a certain point in H_o appears, in image space, to come from a point an equal distance from the axis in H_i. For a thin lens the two planes H_o and H_i coincide with the plane of the lens. With F_o, H_i, H_o, and F_i known, the construction of the image for a given object takes place as before (Fig. 1·25). The two rays from the off-axis point P, through F_o and parallel to the axis, are drawn up to H_o, and the corresponding rays in image space, parallel to the axis and through F_i, respectively, are drawn from the points on H_i with the same axial separation as the points of incidence on H_o. It may be remarked that the distances F_oH_o and

H_iF_i are called the object- and image-side focal lengths. These distances are identical if the indices of refraction or potentials of object and image space are the same. Otherwise, their ratio is equal to the ratio of the corresponding indices or the square roots of the potentials. In view of this and in view of the geometrical construction in Fig. 1·25,

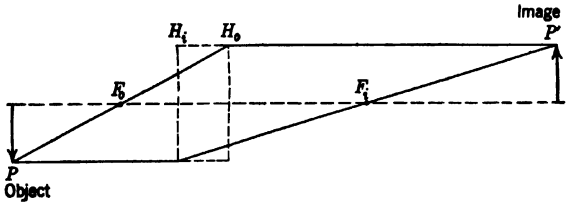


FIG. 1·25. Determination of Image Position and Magnification for a Thick Lens.

the magnification of the image is simply given by the ratio of the image distance (the distance from H_i to the image) and the object distance (the distance from H_o to the object) multiplied by the square root of the ratio of the potentials in object space and image space.

1·4. Motion in Magnetic Fields; Magnetic Electron Lenses. Up to this point the discussion has been confined to electric fields, except for a reference to the use of magnetic fields for deflecting electrons. Magnetic fields with axial symmetry may also be used as electron lenses. It is even possible to define an index of refraction which will govern the motion of electrons in magnetic fields in accord with the laws of ray optics. However, this index becomes dependent on the direction of motion of the electrons. Even though this takes place also in the case of the propagation of light in certain crystals, there is no close analogy between the motion of electrons in magnetic fields and the paths of light rays in any known medium.

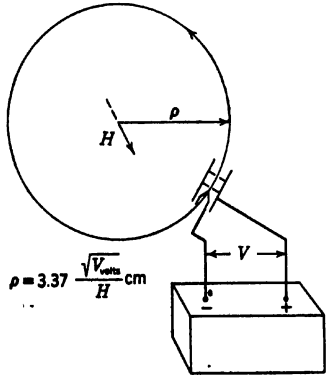


FIG. 1·26. Path of an Electron in a Uniform Magnetic Field at Right Angles to Its Direction of Motion.

A magnetic field exerts a force on an electron at right angles both to the field itself and to the direction of motion of the electron. This force is proportional to the field strength and to the velocity of the electron.

Thus a magnetic field parallel to the motion of an electron exerts no deflecting force on it. On the other hand, a uniform field at right angles to the initial velocity of the electron causes the latter to describe a circle with a radius proportional to its velocity and inversely proportional to the field strength; for a 1-volt electron in a field of 1 gauss the radius is 3.37 centimeters (Fig. 1·26). The frequency of rotation of the electron depends solely on the magnetic field and is proportional to the latter, being 2.8 megacycles per second for 1 gauss.

If, now, the electron is given an added velocity parallel to the field, its path will be extended into a helix, whose pitch is determined by the ratio of this velocity and the field strength (Fig. 1·27). This is the

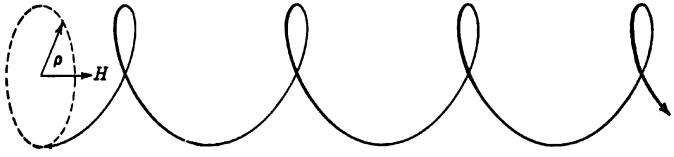


FIG. 1·27. Helical Path of an Electron in a Uniform Magnetic Field.

condition prevailing when a beam of electrons travels in the direction of the magnetic field. Tangential velocity components of the individual electrons will cause them to describe helices, whose several diameters are determined by the magnitude of these components. However, if the axial velocities of the electrons are the same, the pitch of their helices is identical, and all electrons passing through a given point will, after one revolution, come together in a second point, a distance equal to the pitch from the first point in the direction of the magnetic field (Fig. 1·28). In this manner the magnetic field will produce a succession

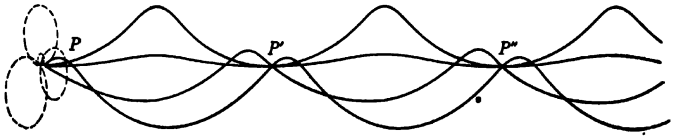


FIG. 1·28. Focusing Effect of a Uniform Magnetic Field.

of equally spaced, identical, erect electron images of an object placed in their path. These images can be rendered visible by placing a fluorescent screen where they occur (Fig. 1·29).

The long magnetic lens formed by a uniform field constitutes an exceptional case and is limited to the forming of images of unity mag-

nification. Another limiting case of greater importance in electron optics is that of the very short magnetic electron lens, with a strong field confined to a small region of space, such as may be produced by

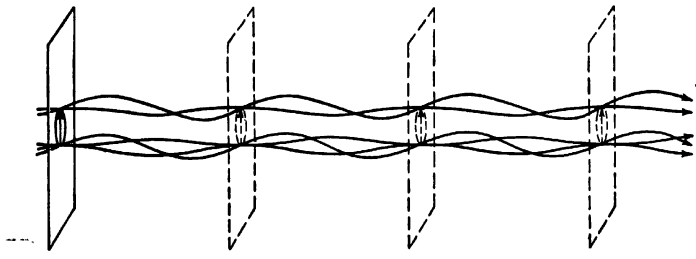


FIG. 1·29. Image Formation by a Uniform Magnetic Field.

iron-encased current-carrying coils with a narrow slit in the iron shell (Fig. 1·30). The magnetic equipotential surfaces, which are related to the magnetic field lines and field strengths as are electric equipotential surfaces to electric field lines and field strengths, show a distribution similar to that of electric equipotentials for a configuration of two electrodes of the same shape as the pole pieces. Thus, also, the magnetic field component normal to the axis is proportional, for small distances from the axis, to this latter distance. For an electron reaching the magnetic field parallel to the axis this vertical field component will cause an acceleration in a direction out of the plane of the figure, normal to the direction of motion of the electron and to the vertical field component. The resulting tangential component of motion is, in turn, normal to the axial component of the magnetic field. Accordingly a force normal to both must act on the electron, causing a deflection of the electron toward the axis; in the second half of the symmetric magnetic field considered, the vertical component of the field, and hence the tangential acceleration, are reversed, but the acceleration toward the axis continues until the electron leaves the field.

Although the tangential velocity component is thus nullified, there remain a displacement of the ray in a tangential direction and an angular deflection of the ray toward the axis by an amount proportional to the original distance of the ray from the axis. Figure 1·31, *a* and *b*, shows the projection of the path on a cylinder about the axis and on a plane through the axis, respectively. It follows that all rays incident parallel to the axis reunite at a point on the axis after having been rotated about the latter by a fixed amount. The deflection toward the axis, or refractive power of the lens, is proportional, as follows from the mechanism

described, to the square of the ratio of the magnetic field strength and the velocity of the electrons, whereas the angular shift is proportional to this ratio.

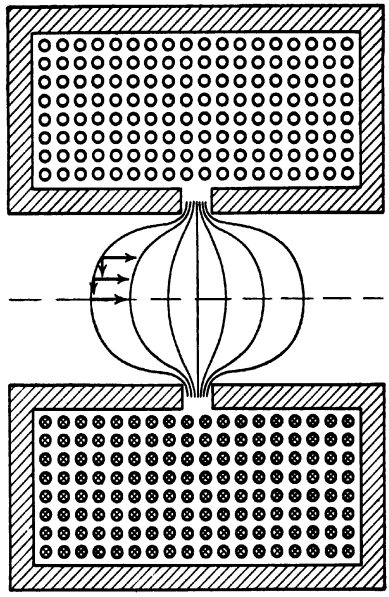


FIG. 1·30. Short Magnetic Lens Formed by an Armored Coil.

Like an electric lens, a magnetic lens, whether short, long, or intermediate, is capable of producing electron images. Their position and

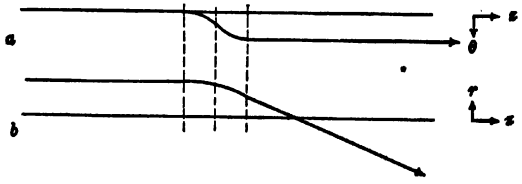


FIG. 1·31. Projection of an Electron Path in a Short Magnetic Lens. (*a*) On a Cylindrical Surface about the Axis. (*b*) On a Plane through the Axis.

magnification may be found if the principal planes and focal points are known. However, these images need not be either erect or inverted,

but may have any orientation as determined by the field distribution. This rotation vanishes only in the case of the extremely thin lens and a lens suitably compounded of magnetic fields acting in opposite direc-

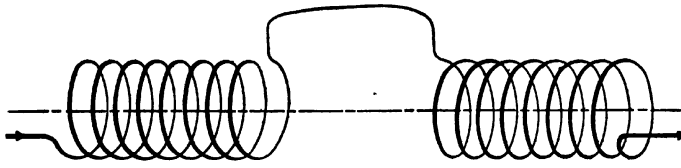


FIG. 1·32. Rotation-Free Magnetic Lens.

tions (Fig. 1·32). Figure 1·33 indicates the action of a magnetic lens of the last type on an electron ray by showing, once more, its projection on a cylinder about the axis (*a*) and on a plane through the axis (*b*).

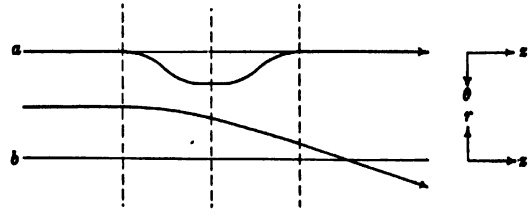


FIG. 1·33. Projection of an Electron Path in a Rotation-Free Magnetic Lens. (*a*) On a Cylindrical Surface about the Axis. (*b*) On a Plane through the Axis.

1·5. General Properties of Electron Lenses; Electron Mirrors. Both magnetic and electric electron lenses — and, for that matter, electron lenses formed by superposed electric and magnetic fields — have certain restrictive properties, owing to the fact that their "lens surfaces" cannot be shaped arbitrarily, as can those of glass lenses. One of the results of this circumstance is that the principal planes H_o and H_i are invariably *crossed*, as shown in Fig. 1·25. The image-side principal plane is closer to a real object than the object-side principal plane. Furthermore, there are no negative electron lenses (lenses which increase the divergence of a ray pencil) which have field-free object and image space. Cathode lenses, that is, lenses in which the electrons leave the object with practically zero velocity, do not fulfil this condition. Figure 1·34 shows, as an example, a negative cathode lens consisting of an anode with a circular aperture in front of a plane cathode. Even here, there is a reduction rather than an increase in divergence if the entire electric field is regarded as a lens.

In addition to electron lenses, there are also electron mirrors. They are obtained if some of the electrodes of an axially symmetric configuration are made so negative that a zero-equipotential surface crosses the

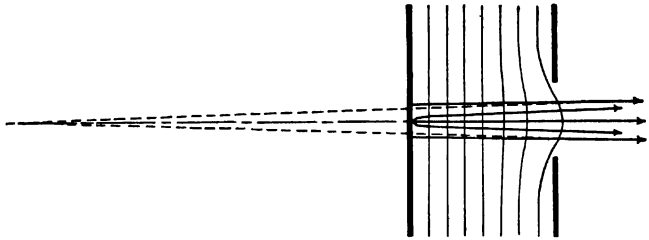


FIG. 1·34. Negative Lens Action of an Anode Aperture in Front of a Plane Cathode.

axis. Figure 1·35 shows an example. Electrons incident on the axis are slowed down as they penetrate into the mirror field until they reach the zero equipotential; at this point their velocity vanishes and their direction of motion is reversed. Electrons striking the mirror at an angle with the axis reverse their direction before this point, as they

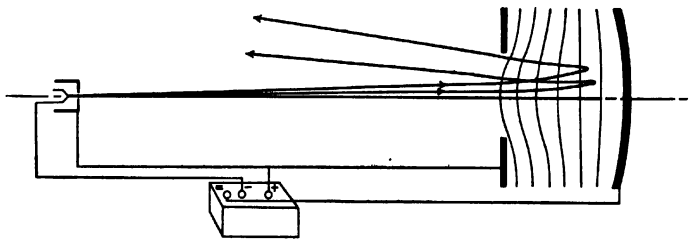


FIG. 1·35. "Convex" Electron Mirror.

retain an appreciable tangential velocity, describing a path resembling a parabola close to the point of reversal. Reflection does not, accordingly, take place at a single surface, as for a light-optical mirror, but is distributed through a region. Nevertheless, a mirror of the type shown in Fig. 1·35 acts optically just as a convex light mirror (Fig. 1·36), forming an erect, virtual image, as is shown in Fig. 1·37. A different arrangement of electrodes and potentials, resulting in a more highly curved zero equipotential, can lead to the analogue of a concave or convergent light mirror (see, for example, Fig. 1·39*e*).

For the electron mirror the conditions of symmetry are such that the object and image focal points and the object and image principal planes,

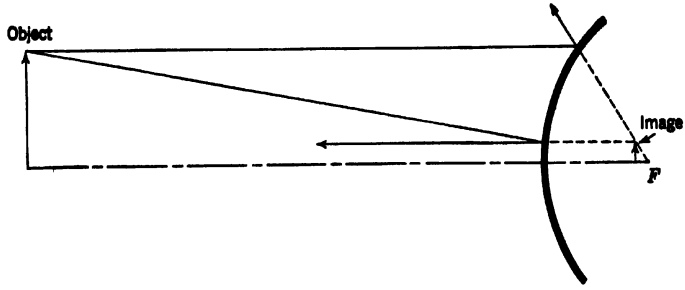


FIG. 1·36. Image Formation by a Convex Light Mirror.

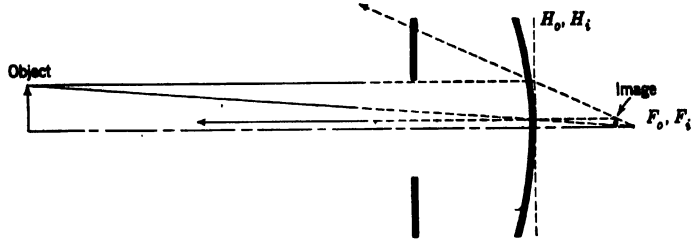


FIG. 1·37. Image Formation by an Electron Mirror.

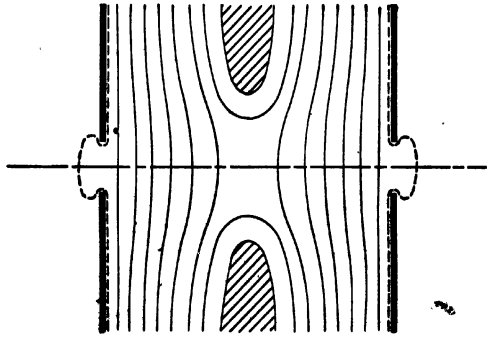


FIG. 1·38. Electric Unipotential Lens.

respectively, must of necessity coincide. Accordingly, also, as for any unipotential lens, or lens bounded by regions of equal potential, the two focal lengths are the same.

Figure 1·38 represents a typical unipotential electron lens formed by three aperture electrodes, of which the outer two are at the same potential. Other simple types of electron lenses and their action on electrons

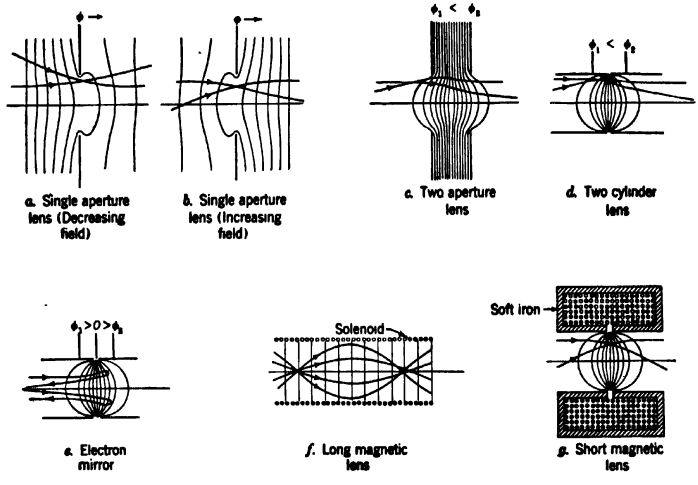


FIG. 1·39. Electron Paths in Various Types of Electron Lenses. (Courtesy *J. Applied Phys.*, Vol. 10, p. 472, 1939.)

passing through them are sketched in Fig. 1·39. In the case of the magnetic lenses the paths indicated represent simply the variation of the distance of the electron from the axis, without regard to the simultaneous rotation about the axis.

1·6. Image Defects of Electron Lenses. So far nothing has been said regarding the quality of the images obtained with electron lenses. A perfect image, disregarding the diffraction effects mentioned in section 1·2, is obtained only if the extent of the object and image as well as the inclination of the image-forming electron rays are exceedingly small, if the velocity of the electrons is uniform, and if the electron concentration at all points of the path is so small that the mutual repulsion of the electrons has negligible effect. Nonfulfilment of these three conditions gives rise to geometrical aberrations, chromatic aberrations, and space-charge defects. A final, obvious, source of image imperfections is mechanical misalignment or defective construction of the electrodes, pole pieces, and coils producing the lens fields.

Both geometric and chromatic aberrations are familiar terms in light optics. If only moderate ray inclinations and image dimensions are considered, it is possible to separate the total effect of the former into five separate effects, which together determine the deviation of the actual image from the ideal, sharp, geometrically faithful image. As in the terminology of light optics, these are referred to as spherical aberration (once regarded as a consequence of the spherical shape of glass-lens surfaces) or aperture defect, coma, curvature of field, astigmatism, and distortion.

Spherical aberration or the aperture defect is the only geometric aberration which causes unsharpness of the image on the optic axis. This aberration results in a circle of confusion about every image point, which is proportional in diameter to the cube of the aperture α of the imaging pencil. This aperture α (Fig. 1·40) is defined as half the vertex

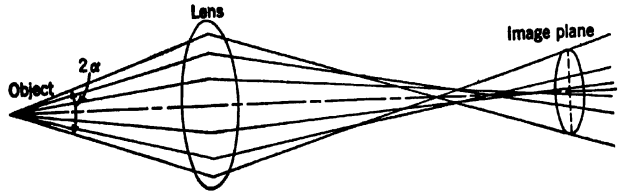


FIG. 1·40. Origin of Spherical Aberration or Aperture Defect.

angle of the ray cone which leaves an object point and is transmitted by the lens. In magnitude the spherical aberration is the same over the entire image. It is caused by the fact that rays passing through the outer parts of a lens are refracted somewhat more strongly than in proportion to the separation from the axis, returning to the latter at a smaller distance from the lens than rays having a smaller inclination to the axis (Fig. 1·40).

The second aberration, coma, derives its name from the fact that it causes the image of a point just off the axis to assume a cometlike appearance. The dimensions of the *comet* are proportional to the square of the aperture of the imaging pencil and to the first power of the separation of the object point from the axis. The presence of coma corresponds to a variation in the magnification of the image with the aperture of the imaging pencils (Fig. 1·41). Although for purely electric, as well as for light-optical, lenses the comet is always directed toward or away from the axis, it may have any orientation for a magnetic lens.

The third aberration is curvature of field. If curvature of field alone is present, a perfectly sharp image may be obtained on either a concave

or a convex image surface. It arises if the rays adjoining the principal rays from the several off-axis object points converge on the latter ahead of or behind the image plane (Fig. 1·42).

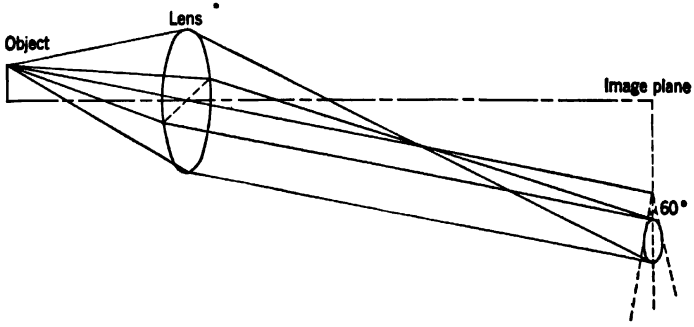


FIG. 1·41. Origin of Coma.

More commonly, such off-axis pencils do not converge at a point, but rather successively in two mutually perpendicular line segments (Fig. 1·43). In this case astigmatism is present. Without simultaneous curvature of field the image of an off-axis object point is a circle of

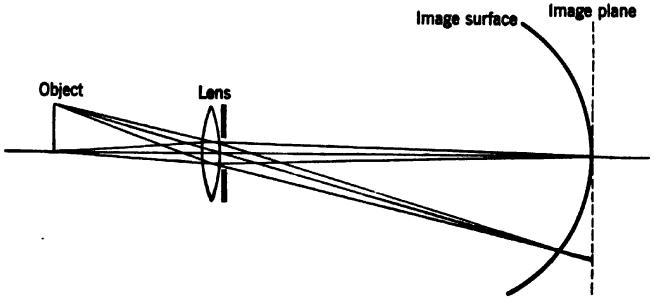


FIG. 1·42. Origin of Curvature of Field.

diameter proportional to the aperture of the imaging pencil and to the square of the distance of the object point from the axis, the image plane being located midway between the *tangential* and *sagittal* image surfaces containing the two sets of mutually perpendicular line segments referred to above (Fig. 1·44). For a purely electric field these line segments are tangential and radial, respectively, accounting for the terms applied to the image surfaces. For magnetic lenses the segments may have any

orientation, so that here the notation loses its significance. In the presence of both astigmatism and curvature the image plane does not fall midway between the two image surfaces, and the image of an off-axis point becomes an ellipse in place of a circle.

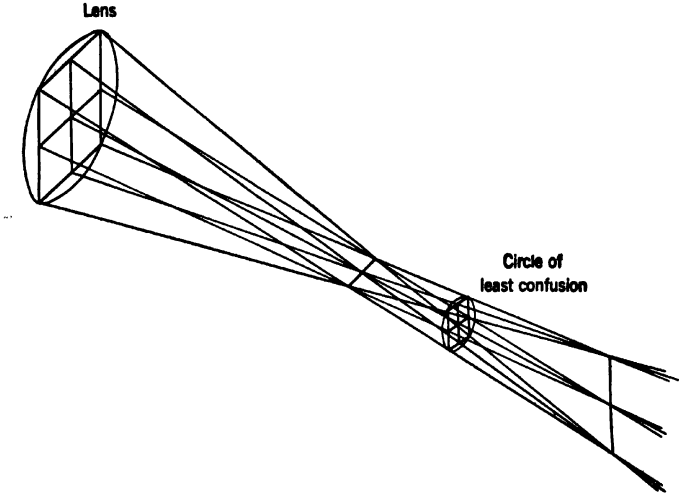


FIG. 1·43. Convergence of a Strongly Astigmatic Ray Pencil Originating at a Single Off-Axis Object Point.

Lastly, there is distortion, an aberration which detracts only from the faithfulness and not from the sharpness of the image. The actual

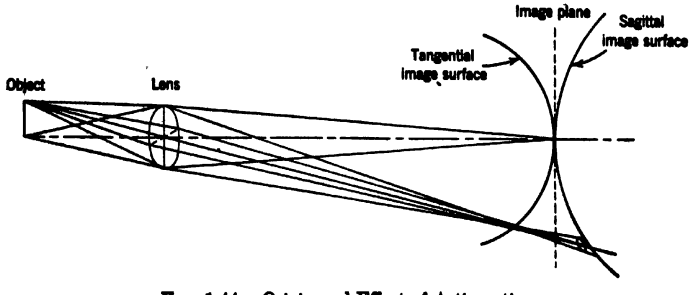


FIG. 1·44. Origin and Effect of Astigmatism.

image point deviates from that in a geometrically faithful image by an amount proportional to the cube of the distance of the object point

from the axis (Fig. 1·45). For purely electric, as for light-optical, lenses this deviation is toward or away from the axis, leading to *barrel* or *pincushion* distortion (Fig. 1·46*a* and *b*), whereas for magnetic lenses a twist, or rotational distortion, may be added to this stretching or contraction (Fig. 1·46*c*).

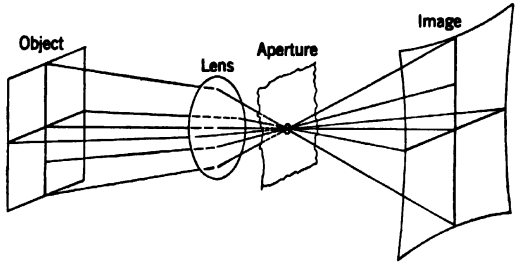


FIG. 1·45. Origin of (Pincushion) Distortion.

Although, normally, all these geometrical aberrations occur together, the conditions of image formation may be such as to single out one or two of them at a time. Thus the image of a point on the axis shows spherical aberration only; that of a small area very close to the axis, imaged with wide-aperture pencils, will show coma as well as spherical

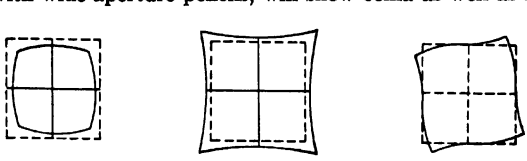


FIG. 1·46. Effect of (*a*) Barrel; (*b*) Pincushion; and (*c*) Rotational Distortion on a Square Pattern.

aberration. An image covering a large field, produced by narrow electron pencils, will be affected principally by curvature of field, astigmatism, and distortion. The same image, formed with pencils of infinitesimal aperture, will show distortion only. Figure 1·47 indicates the effect of the individual aberrations and their sum on the image of a fine grid. The image representation is imperfect insofar as no attempt has been made to represent " grays " as distinct from " blacks " in the image.

The chromatic aberrations in light optics arise from the fact that light of different color (wave length) is refracted to a different degree, leading, for example, in the case of a simple glass lens, to a shorter focal

length for blue light than for red (Fig. 1·48*a*). For electrons the index of refraction is given, at least in part, by the velocity of the electron. Hence variations in the initial velocities of the electrons result in a different refractive power of the electron lenses, thus influencing their

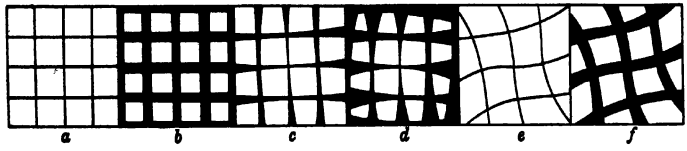


FIG. 1·47. Appearance of the Image of a Square Pattern in the Presence of the Geometrical Aberrations of an Electron Lens. (*a*) Original Pattern. (*b*) Aperture Defect. (*c*) Coma. (*d*) Curvature of Field and Astigmatism. (*e*) Distortion. (*f*) Combination of All Preceding Aberrations.

paths (Fig. 1·48*b*). An increase in initial velocity signifies an addition to the refractive index which is greater at points of low potential than at those of higher potential. Thus the ratio of the refractive indices in adjoining regions and, accordingly, the refractive power of an electron

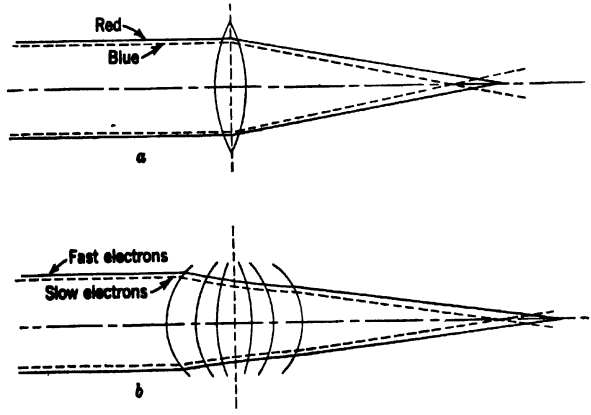


FIG. 1·48. Chromatic Aberration for (*a*) Light and (*b*) Electron Lenses.

lens are reduced. From another point of view, the deflecting fields of the lens act for a shorter time on electrons of greater velocity, leading to the same result.

Usually two types of chromatic aberration are considered separately, chromatic difference in image position and chromatic difference in magnification. For a unipotential thin lens the second vanishes, since the magnification of the image is given simply by the ratio of the image

distance to the object distance from the lens. For a thick lens, on the other hand, both focal points and principal planes may shift in such a manner that the two chromatic aberrations may arise either separately or together. Figure 1·49 shows the condition, fulfilled in many large-aperture light-microscope objectives, of zero difference in image position

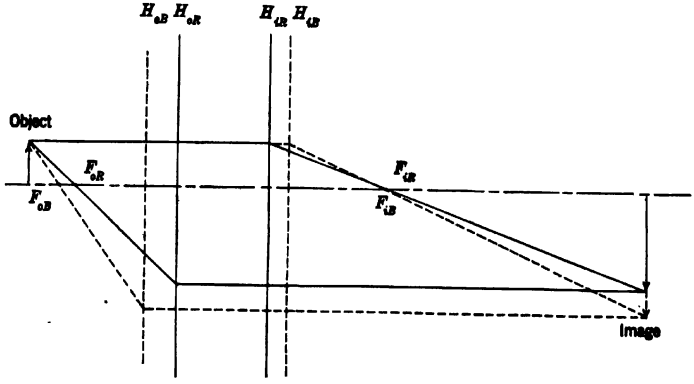


FIG. 1·49. Chromatic Difference in Magnification with Zero Difference in Image Position.

and appreciable difference in magnification. Magnetic electron lenses normally possess, in addition to chromatic difference in magnification, chromatic difference in rotation. The image varies in both size and orientation with the initial velocity of the image-forming electrons.

Space-charge effects, finally, are peculiar to electron optics, as no mutual influence of light rays, such as repulsion or scattering, has been

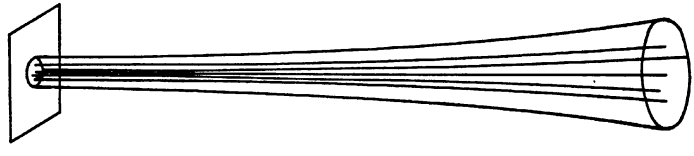


FIG. 1·50. Spreading of Originally Parallel Electron Beam Due to Space Charge.

observed up to the present. A strong concentration of electrons at any point has the effect of reducing the potential or decreasing the speed of the individual electrons. In a beam of given cross section this concentration is inversely proportional to the velocity. Thus space charge plays a role primarily in the case of low-velocity, high-current beams. Here it produces a spreading of the beam, resembling the effect of a negative lens (Fig. 1·50).

One of the primary aims of optical design is the reduction of the aberrations to a minimum. In light optics this has been accomplished to a remarkable extent. Thus light-microscope objectives are made whose aperture defect, chromatic aberration, and coma have been nullified to the extent that an adequate image field, whose sharpness is limited only by the diffraction effects mentioned in section 1·2, may be obtained with imaging pencils with an aperture angle of 72 degrees. Nothing comparable has been achieved in electron optics. Although this may be ascribed in part to the youthfulness of electron optics, it rests unquestionably also on the basic drawback of the impossibility of arbitrarily shaping the lens surfaces, resulting, in part, in the non-existence of negative electron lenses.

This circumstance, at the same time, makes it impossible to construct a practical[1] electron lens free from spherical aberration. The condition of zero aperture defect can be approached only by the device of reducing the dimensions and focal length of an imaging lens (and, hence, of the area imaged) toward zero, a process which is limited by mechanical skill and the properties of materials. A study of favorable lens fields obtainable with present means indicates that the best type of lens may have a spherical aberration about one-sixth as large as that of a simple glass plano-convex lens. Electron lenses in use at the present time are found to have appreciably larger aberration than this.

The situation with regard to the remainder of the geometrical aberrations is less serious. Thus specific lens fields may be prescribed for which coma, distortion, and astigmatism all vanish, even though the fields in question are not readily realized in practice. The effect of curvature of field may be nullified by giving the object a compensating curvature, a procedure which may, in some cases, result in a simultaneous reduction in astigmatism and distortion. Coma and distortion both vanish for a system consisting of completely antisymmetric electric and magnetic fields giving unity magnification.

The elimination of chromatic aberration is subject to restrictions similar to those applying to spherical aberration. For all electron lenses the (real) image plane lies closer to the object for slow electrons than for fast electrons (Fig. 1·51). For a short magnetic lens the percentage change in focal length is given directly by the percentage change in kinetic energy of the electrons. Thus, to obtain no more chromatic aberration with such an electron lens than with a simple lens of crown glass, the kinetic energy of the electrons (or the accelerating voltage) must fluctuate by less than 2 per cent. To render the effect of chro-

[1] This refers to lenses formed by static fields. Regarding variable-field lenses, see section 17·2.

matic aberration insignificant, it is customary to reduce the variations in electron velocity by minimizing the fluctuations in the accelerating voltage.

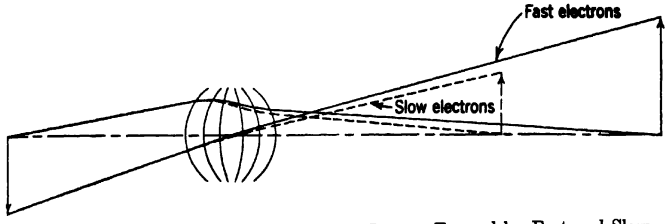


FIG. 1·51. Relative Position of Electron Images Formed by Fast and Slow Electrons.

If the emission velocities of the electrons from a cathode determine the velocity fluctuations or if these are occasioned by energy losses in the object traversed by the electrons, this is not possible. However, it is possible to attain a certain correction of chromatic aberration by combining an electron mirror with an electron lens system (Fig. 1·52). In

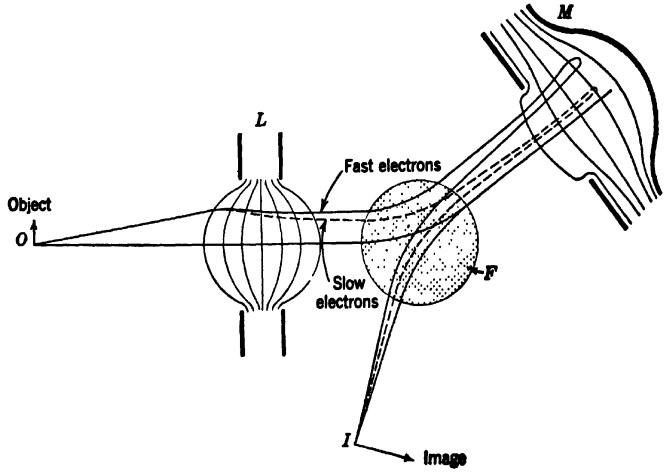


FIG. 1·52. Correction of Chromatic Aberration with the Aid of an Electron Mirror.

the figure an object at O is imaged by a lens L in conjunction with the mirror M at I, a magnetic field normal to the plane of the paper serving to deflect the rays just sufficiently that a fluorescent screen or plate may receive the image without interfering with rays coming from the lens.

The faster electrons, which enter the mirror in a slightly divergent pencil, penetrate deeper into the mirror and are thus subject to a larger fraction of the converging field. As a result it is possible to cause them to be focused at the same point as the originally more convergent, slower electrons.

The electron mirror, similarly, offers a possibility, in principle, of obtaining a correction of the aperture defect. A hindrance to the utilization of this property rests in the practical difficulty of arranging the imaging fields in such a manner that the presence of the object does not influence them or interfere with the reflected electrons without enhancing image defects by the introduction of deflecting fields.

1·7. High-Voltage Electron Optics; Ion Optics. In the considerations up to this point it has been assumed that the electrons forming the image had velocities small compared to the velocity of light. In many important applications of electron optics — in particular, electron microscopy and electron-accelerating devices — this condition is not fulfilled.

For speeds comparable to that of light, the electron velocity ceases to be proportional to the square root of the potential. As the latter is increased, the velocity approaches a constant value, namely, the velocity of light, $3 \cdot 10^{10}$ centimeters per second. The electron index of refraction (in the absence of a magnetic field), on the other hand, increases more rapidly than the square root of the potential, eventually becoming proportional to the potential itself (Fig. 1·53).

The result, for both electric and magnetic lenses, is similar. Thus, for low velocities, the focal length of an electric lens remains unaltered if the potential on all the electrodes (relative to the cathode) is changed in the same ratio. A unipotential lens, such as that shown in Fig. 1·54, may, for low voltages, be operated with an alternating voltage. No electrons leave the cathode during the negative half of the cycle; during the positive half both position and magnification of the image remain constant. As the overall voltage rises to higher values, however, the focal length increases, the relative change in focal length being proportional, to a first approximation, to the overall voltage. For 100,000 volts it amounts to 2 per cent for a particular lens of the type shown in Fig. 1·54.

The same thing is true for a magnetic lens, for which the ratio of the applied voltage to the square of the magnetic field is kept constant, leaving the shape of the field distribution unaltered as well. However, the relative change in focal length is here larger than for the electrostatic lens. For a short lens it is equal — independently of the field distribution of the magnetic lens — to approximately one-millionth of

the applied voltage. Accordingly, for 100,000 volts the deviation in focal length is one-tenth (or 10 per cent) of the focal length itself.

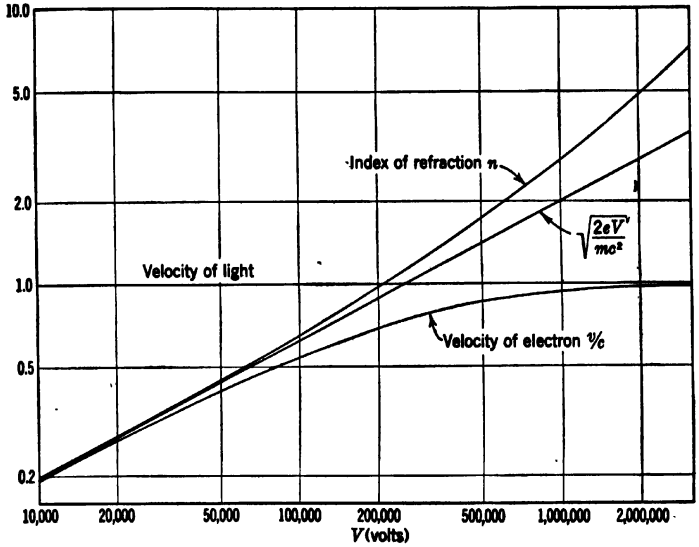


FIG. 1·53. Electron Index of Refraction and Electron Velocity for High Accelerating Voltages.

The aberrations are not changed in character by passing over to very high accelerating potentials and electron velocities, though their relative magnitudes may be affected.

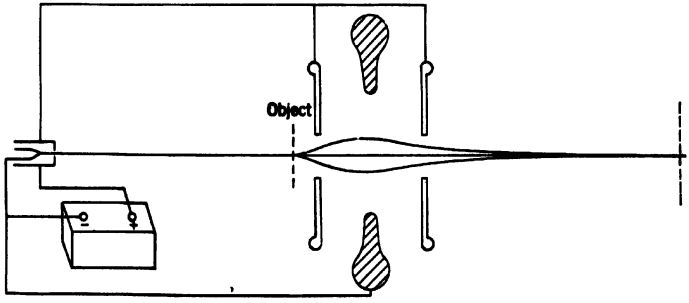


FIG. 1·54. Unipotential Lens Operated with Single Voltage above Ground.

In closing, reference should be made to the use of charged particles other than electrons, for example, atomic ions, for the formation of

images. The laws governing their motion are the same as for electrons, all differences arising merely from their difference in mass. The lightest of the ions, the proton or positively charged hydrogen atom, has a mass approximately 1840 times as great as that of the electron. This results in:

1. A shorter wave length, for given accelerating voltage, in inverse proportion to the square root of the masses, signifying a factor of 43 for the proton, and correspondingly reduced diffraction effects.
2. A smaller velocity, for given accelerating voltage, in inverse proportion to the square root of the masses, resulting in a greater focal length for magnetic lenses in direct proportion to the masses, signifying a factor of 1840 for the proton.

The focal length of an electric lens, on the other hand, does not depend on the mass of the particles. This makes electric lenses the natural imaging means for high-voltage ions.

If ions, in spite of their smaller diffraction effects, have not been used to any extent for imaging purposes, this is, without doubt, due primarily to the greater difficulty of generating ion beams, especially intense, narrow, homogeneous ion beams of very small divergence. Enhanced space-charge effects (due to the smaller velocity) and reduced penetration of matter are added drawbacks.

CHAPTER 2

APPLICATIONS OF ELECTRON OPTICS

2·1. Electron Guns. The first application of electron-optical principles to practical problems was concerned with the formation of narrow electron beams of high intensity in electronic oscillograph tubes. A small bright spot is formed at the incidence of the beam on a fluorescent screen. This is displaced when a difference of potential is applied to a pair of deflecting plates or a current flows through deflecting coils, subjecting the beam to a transverse electric or magnetic field (Fig. 2·1).

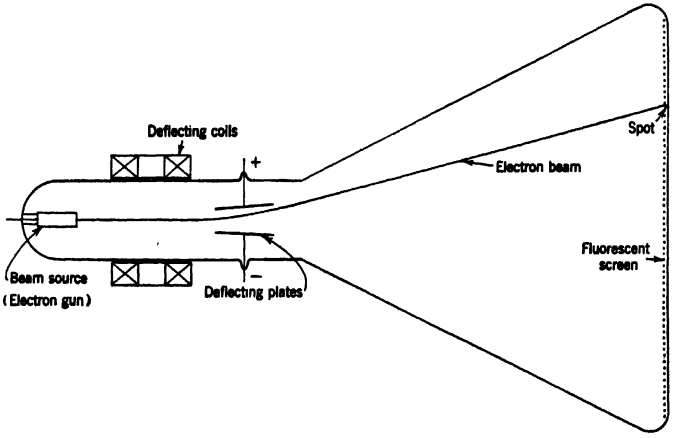


FIG. 2·1. Schematic Diagram of Oscillograph Tube with Electric and Magnetic Deflection.

The displacement of the spot can thus be used to measure either voltages or currents. Similar needs, involving ever-smaller spot sizes and greater beam intensities, arose in connection with the rapidly developing technique of electronic television. Here an electrically focused scanning beam is used to convert an optical image into a sequence of electric signals in the pick-up tube or iconoscope,[1] and a second similar beam to reconstruct the image on the viewing screen of the receiver or kinescope[2]

[1] See Zworykin, reference 1.

[2] See Zworykin, reference 2.

(Fig. 2·2). The beam-forming system itself, comprising a cathode and one or more electron lenses, has come to be known as an *electron gun.*

The simplest electron gun consists of a small cathode and one electron lens, the electron spot being the image of the cathode (Fig. 2·3). To obtain a small spot, the lens must be placed at a considerable distance

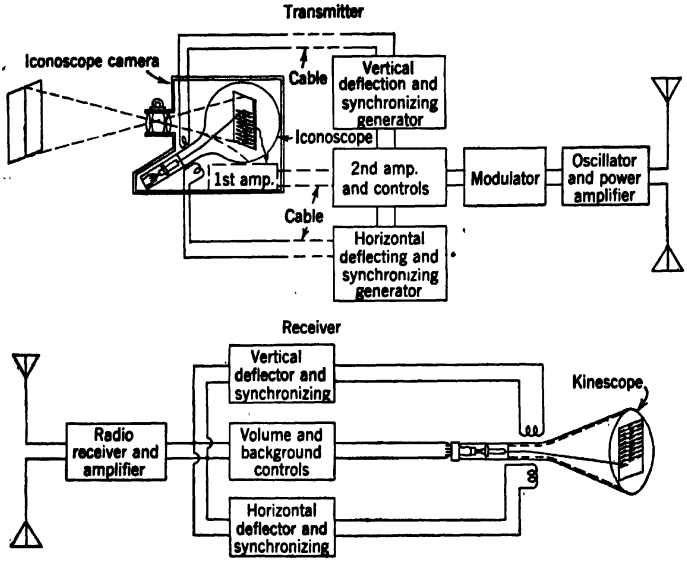


FIG. 2·2. Schematic Diagram of an Electronic Television System. (Courtesy *Proc. Inst. Radio Engrs.*, reference 1.)

from the cathode, the magnification being proportional to the ratio b/a of the image and object distances if the electrons are accelerated to their full velocity close to the cathode. If a uniform accelerating field is

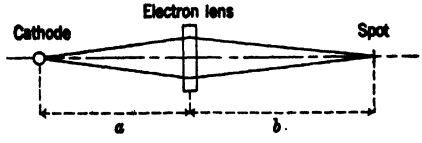


FIG. 2·3. The Simplest Type of Electron Gun

applied between the cathode and the lens, the magnification will be just $b/(2a)$, since the accelerating field has the effect of causing the object, as seen from the lens, to appear to lie at twice the actual distance

from the lens. In this case the random initial velocities of the electrons will cause them to spread out into a relatively broad beam in the weak accelerating field preceding the lens. Consequently, a large portion of the lens is involved in the imaging process, and the lens aberrations will result in a spot size larger than that expected from the known magnifica-

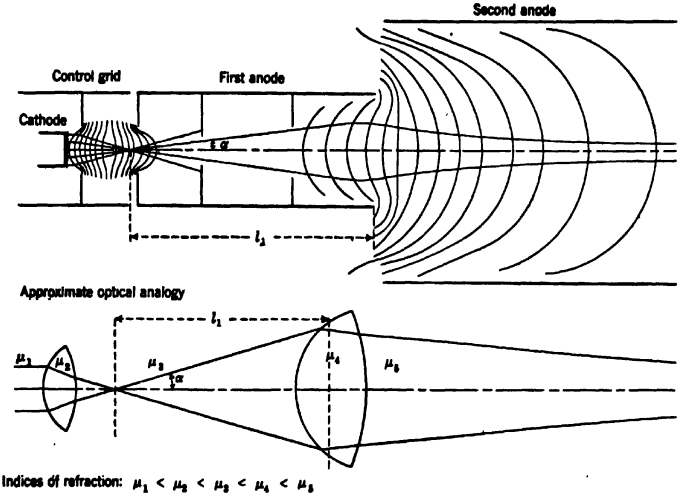


FIG. 2·4. Typical Electrostatic Electron Gun and Light-Optical Analogue.

tion ratio. This simple gun has no provision for modulating the intensity of the spot. Moreover, it is incapable of furnishing large intensities since, even for moderate beam current, the negative field of the electrons in front of the cathode causes electrons to return to the cathode.

For all these reasons the somewhat more complex two-lens configuration shown in Fig. 2·4[3] has quite generally replaced the single-lens gun. This arrangement, with either the first or both lenses electrostatic, may be regarded as the prototype of nearly all modern electron guns. The outward appearance of such a gun — more specifically, of the type represented in Fig. 2·4 — is shown in Fig. 2·5, which presents a kinescope gun. The second anode is missing, as it is formed by a conducting coating on the wall of the tube in which the gun is mounted.

As seen most clearly in the light-optical analogue (Fig. 2·4), the principal rays, formed by the electrons emitted by the cathode with zero initial velocity, proceed normal to the surface of the cathode in the

[3] See Zworykin, reference 3.

accelerating field immediately in front of the latter. They then enter the first lens formed between the cathode, the control grid, and the

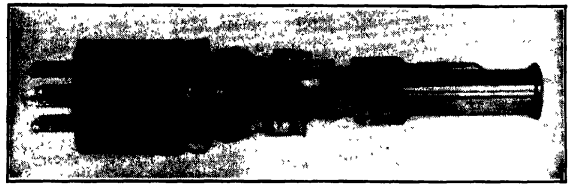


FIG. 2·5. Electron Gun of a Kinescope. (Zworykin and Morton, *Television*, John Wiley and Sons, New York, 1940.)

first-anode aperture. This lens converges the principal rays to a point a small distance beyond it, the *crossover*. The second lens, formed in the example shown by the accelerating field between the first and

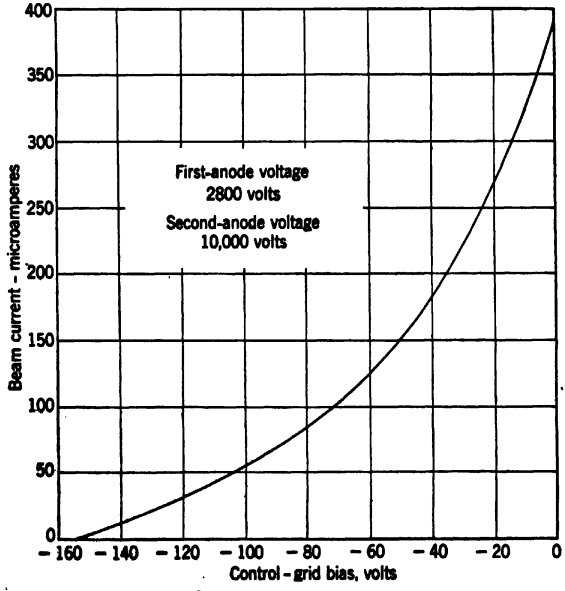


FIG. 2·6. Variation of Beam Current with Grid Bias for an Electrostatic Projection Tube Gun. (Zworykin and Painter.)

second anodes (two coaxial cylinders of different diameters), images the spot on the screen. Thus the spot is not the image of the cathode, but of the crossover.

The control grid is usually maintained at a potential slightly negative

with respect to the cathode; it is made more negative if the beam current is to be reduced. Only a certain central zone of the grid aperture will permit the passage of electrons with a given initial velocity. Those incident on the marginal zones will be turned back toward the cathode. With increasingly negative grid potential the diameter of the central zone is reduced. Thus changing the grid voltage is roughly equivalent to changing the size of the grid aperture. The actual rate of change of the effective size of the aperture with the grid potential, for a gun of the type shown in Fig. 2·4, is plotted in Fig. 2·6.[4] Although the grid potential influences, to a certain extent, the strength of the first lens and hence the position of the crossover and the final spot size, this effect is normally of minor importance and can largely be eliminated by inserting additional apertures at fixed potentials near the control grid.

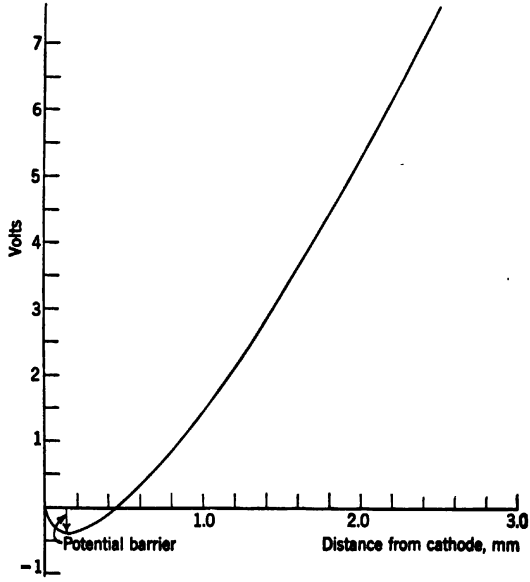


FIG. 2·7. Potential Barrier in Front of a Cathode Formed by Space Charge.

In high-current guns the beam-suppressing action of the grid is largely supported by the space charge of the beam itself. The negative charge of the relatively slow electrons in front of the cathode forms a potential barrier which causes electrons leaving the cathode with low initial velocities to return to it (Fig. 2·7). Thus, in special cases,

[4] See Zworykin and Painter, reference 4.

adequate control may be exercised with a control grid maintained at a positive voltage. An example of this is a projection-tube gun designed by Law (Fig. 2·8)[5] which employs a magnetic second lens.

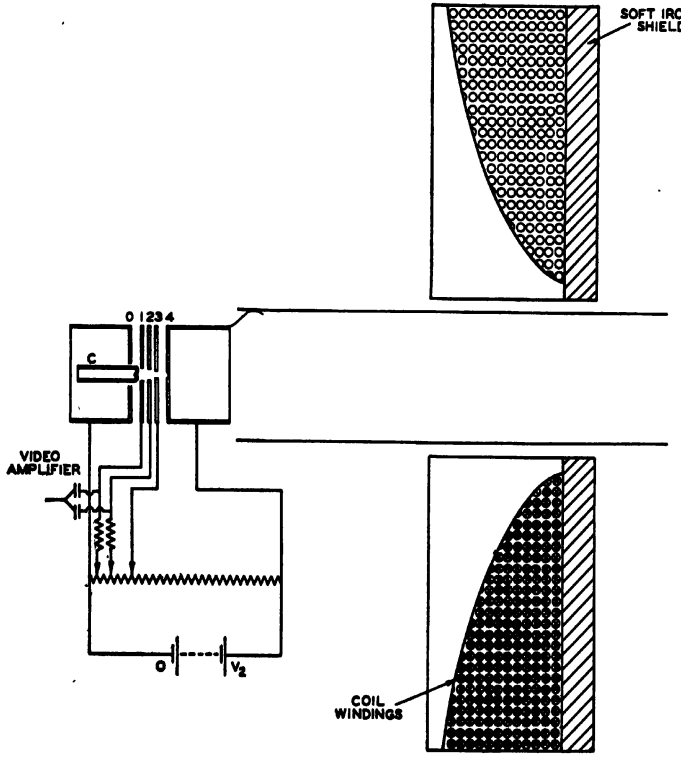


FIG. 2·8. Projection Gun with Magnetic Second Lens. (Law.) (Courtesy *J. Applied Phys.*, Vol. 10, p. 477, 1939.)

In electrostatic guns of the type shown in Figs. 2·4 and 2·5 the first anode normally is maintained at about one-fourth of the voltage applied to the second anode. This ratio is fixed by the following consideration: As shown in Fig. 2·4, the diameter of the beam entering the second lens is given by $2l_1\alpha$. Here l_1 is the distance between the crossover and the second lens. It may, for relatively large gun-to-screen distances, be approximated by F_1, the object-side focal distance of the second lens. α, the angle of divergence of the principal rays leaving the crossover,

[5] See reference 5.

on the other hand, depends exclusively,[6] for a given construction, on the cathode area utilized. This follows from the fact that the paths of the principal rays, that is, the rays leaving the cathode with zero initial velocity, in the first-lens region are independent of the magnitude of the first-anode voltage.

If the diameter of the second-lens area utilized by the beam, $2l_1\alpha \cong 2F_1\alpha$, is greater than a certain fraction k of the diameter of the first anode — the *gun diameter* D_1 — the aberrations of the second lens will cause undue spreading of the final spot. It is thus necessary to demand that

$$2F_1\alpha < kD_1 \qquad [2\cdot1]$$

The value of k may be derived from Fig. 2·9. The latter shows measurements by Epstein[7] of the final spot diameters obtained with various second lenses of the character considered, as function of the fraction of the gun aperture utilized. It is seen that for $k > 0.30$ the spreading of the spot with an increase in the beam diameter is rapid. Equation 2·1 thus imposes a definite upper limit for F_1/D_1, that is, the object-side focal distance in gun diameters, provided that α, or the magnitude of the cathode surface utilized, is prescribed. Since, in turn, as shown in Fig. 2·10[8] applying to the specific case of a ratio 2.3 between the second- and first-anode diameters, F_1 decreases as the voltage ratio is increased, Equation 2·1 also yields a lower limit for the voltage ratio of the second lens. This usually falls, substituting reasonable values for α, in the neighborhood of 4.[9]

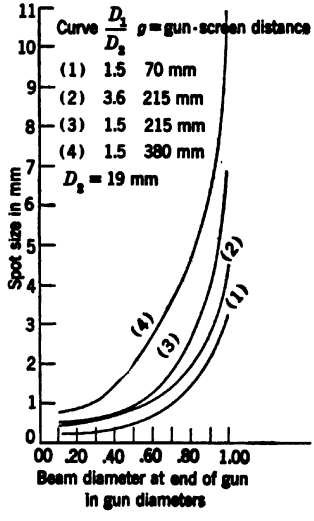


FIG. 2·9. Variation of Spot Diameter with Width of Beam at Second Lens. (Epstein.) (Courtesy *Proc. Inst. Radio Engrs.*, reference 6.)

In the presence of space-charge limitation of the beam current, the attainment of the maximum intensity in a spot of prescribed size makes

[6] Neglecting space charge.

[7] See reference 6.

[8] See Epstein, reference 6.

[9] It is obvious that in the above elementary treatment a number of factors entering into the choice of the gun dimensions have been omitted from consideration. The conclusion is, however, substantially correct.

it desirable to apply a strong electrostatic field in the immediate neighborhood of the cathode.[10] This may be accomplished by the

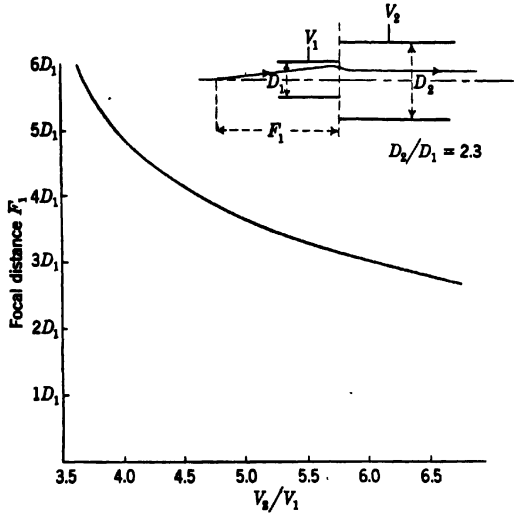


FIG. 2·10. Variation of Focal Distance with Voltage Ratio for Lens Formed by Two Coaxial Cylinders with a Diameter Ratio 2.3. (Epstein.)

insertion of an accelerating grid between the control grid and the first anode, which, simultaneously, reduces the dependence of the spot size on

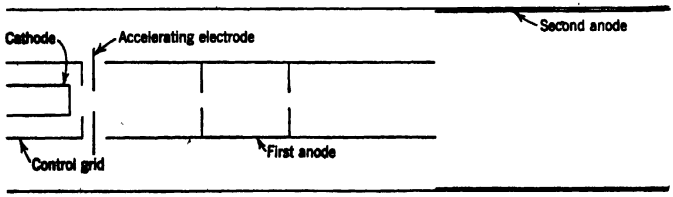


FIG. 2·11. Electron Gun Employing an Accelerating Electrode between the Control Grid and the First Anode.

the grid bias[11] (Fig. 2·11). Alternatively, the full anode voltage may be applied to the first anode and the second lens may be made a unipo-

[10] The emission current from the cathode increases as the $\frac{3}{2}$ power of the field at its surface, i.e., for a given geometrical arrangement,

$$I_e = cV_1^{3/2} \qquad [2\cdot2]$$

where V_1 is the potential of the nearest accelerating electrode.

[11] See Zworykin, reference 7.

tential lens. An example of such an arrangement is Law's projection tube (Fig. 2·8), which employs a magnetic second lens. Iams[12] and Janes and Hickok[13] have proposed electrostatic unipotential lenses with negative center element to accomplish the same purpose. The spot size itself depends, to a first approximation, only on the final energy of the electrons forming the spot — the *overall* applied voltage — and not on the distribution of voltage between the first and second anodes. Its area is proportional to the ratio of the absolute temperature of the cathode to the final kinetic energy of the electrons.[14] The maximum current density attainable in the spot, on the other hand, is proportional to the reciprocal of this ratio, to the emission current per unit area of

[12] See reference 8.

[13] See reference 9.

[14] This may readily be demonstrated. Electrons leaving a given point of the cathode deviate from the principal ray by a distance r which is proportional to the ratio of their initial radial velocity to the velocity which they acquire in traveling to the point in question:

$$r = f(z)\cdot\frac{v_r}{v_z} = g(z)\left(\frac{V_r}{V_1}\right)^{1/2} \qquad [2\cdot3]$$

Here $mv_r^2/2 = eV_r$ and $f(z)$ and $g(z)$ are some functions of the axial distance z from the cathode. For the region just in front of the cathode this is obvious, since the lateral displacement is simply given by $v_r t$, and the time t is given by $v_z = -eEt/m = \text{const } V_1 t$. Here E is the field in front of the cathode and

$$\frac{mv_z^2}{2} = eV = -eEz = \text{const } V_1 z$$

For the remainder of the path Eq. 2·3 is vouchsafed for rays passing close to the paraxial principal rays by the proportionality of all radial field forces to the radial separations (Figs. 1·19 and 1·20). Hence the diameter of the crossover, which is practically identical for all the principal rays, is given by

$$d_c = C\cdot\left(\frac{V_r}{V_1}\right)^{1/2} \qquad [2\cdot4]$$

If the effects of aberrations are neglected, the diameter d_s of the spot is equal to the product of d_c and the magnification. The latter is given, as in light optics, by the ratio of the image and object distances, b (lens to screen) and a (crossover to lens l_1), multiplied by the ratio of the indices of refraction in object and image space, $n_1/n_2 = (V_1/V_2)^{1/2}$. Thus

$$d_s = C\left(\frac{V_r}{V_1}\right)^{1/2}\left(\frac{V_1}{V_2}\right)^{1/2}\frac{b}{a} = \text{const}\left(\frac{V_r}{V_2}\right)^{1/2} \qquad [2\cdot5]$$

For a thermionic cathode the mean energy associated with lateral motion at emission is kT, where k is Boltzmann's constant ($1.37\cdot10^{-16}$ erg/degree) and T is the temperature of the cathode measured on the absolute scale. This energy is the same as that given an electron in being accelerated through a potential difference of $T/11{,}600$ volts. The area of the spot, $\pi d_s^2/4$ is, accordingly, proportional to T/V_2. It is independent of the first-anode voltage.

the cathode, and to the square of the convergence angle of the beam at the spot.[15]

In conclusion, it should be mentioned that electron guns employing a magnetic second lens, that is, the condenser, are also utilized in the electron microscope (Fig. 2·12). However, since, normally, the condenser lens is quite separate from the remainder of the gun, that is, the source, anode, and, eventually, grid-control unit, it has become customary to refer to the latter assembly without the condenser as the

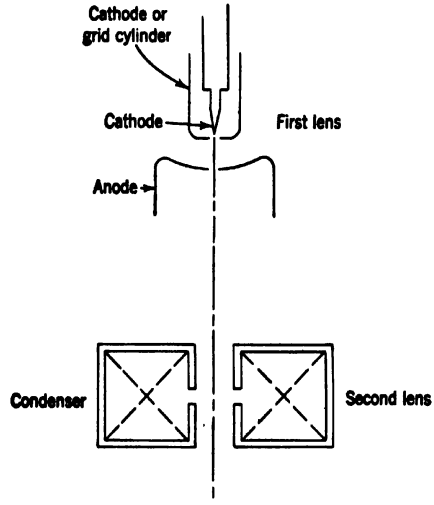


FIG. 2·12. Beam-Forming System of Electron Microscope as Electron Gun (Schematic).

electron gun of the microscope. Furthermore, the function of the electron gun in the electron microscope is not to form a spot of maximum brilliance and minimum dimensions, but rather to illuminate the specimen with electron rays with a divergence or convergence appropriate for the purpose at hand.

2·2. Image-Reproducing Devices. However great the practical importance of the electron gun may be, it represents only an imperfect demonstration of the true optical character of axially symmetric electric and magnetic fields. To the layman, the most characteristic property of optical systems is the capacity of forming faithful images of given objects. A lens is thought of more commonly as an image-reproducing device than as a light- (or electron-) concentrating instrument.

[15] See section 10·5.

The principal image-reproducing devices of light optics are the microscope, the telescope, and the camera (including the picture projector as a modification). In electron optics the principal image-reproducing devices are the electron microscope and the image tube (including the image multiplier as a modification). The electron microscope will be the subject of the succeeding seven chapters, so that only the image tube need be discussed here. Electron telescopes in the true sense do not exist, since the observation of objects at a great distance by means of electrons emitted or transmitted by them would require the creation of an evacuated path between the distant object and the observer. The term *electron telescope* has been applied to an instrument combining an image tube with light-optical lenses in such fashion as to make possible the observation of objects at a distance by means of radiations to which the eye is insensitive.[16]

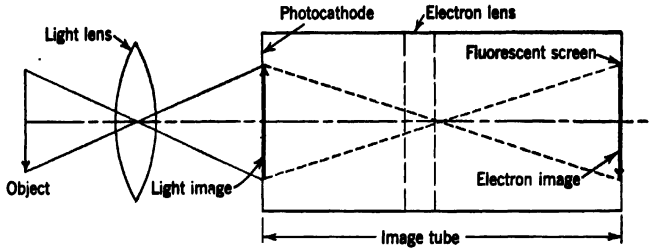


FIG. 2·13. Schematic Diagram of an Image Tube.

The image tube (Fig. 2·13) corresponds most closely to the camera, forming an image of an electron emitter, most commonly a photocathode on which a light image is projected, with a relatively low magnification or reduction factor. This image is observed on a luminescent screen. The image tube differs from the camera in the possibility of greatly increasing the energy of the imaging medium, that is, the electrons leaving the cathode, in the course of the formation of the image by applying a high voltage between the cathode and the viewing screen.

Early applications of the image tube[17] were largely concerned with utilizing the sensitivity of photosensitive surfaces to invisible, ultraviolet or infrared, radiations for converting invisible light images into visible images that can be viewed directly on a fluorescent screen. Figure 2·14*a* shows a particular image tube with this function. It consists simply of a semitransparent photocathode and a closely spaced

[16] See Zworykin and Morton, reference 10.

[17] See Holst, de Boer, Teves, and Veenemans, reference 11, and Zworykin and Morton, reference 10.

screen at a high potential without any focusing device. Excessive spreading of the photoelectrons emitted by any point of the cathode is prevented by the great strength of the field at the surface of the cathode, together with the shortness of the path of the electrons. Much

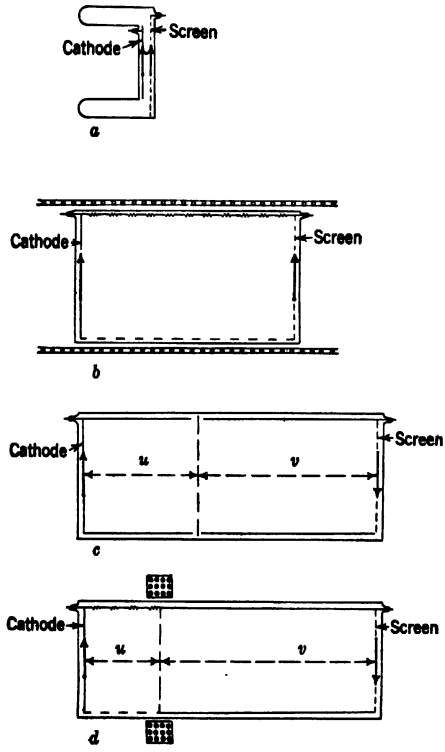


FIG. 2·14. Various Types of Image Tubes. (*a*) Simple Electric Accelerating Field. (*b*) Superposed Uniform Electric and Magnetic Fields. (*c*) Electrostatic Lens. (*d*) Short Magnetic Lens.

sharper pictures can be obtained, however, by employing a focusing field. In Fig. 2·14*b* a uniform magnetic field formed by a solenoid surrounding the image tube has been superposed on the electric accelerating field. Because of their initial tangential velocities the electrons describe helices with their axes parallel to the field direction and with a frequency of revolution $\nu = eH/(2\pi mc)$, H being the strength of the magnetic field and c the velocity of light. Focusing will take place if

the length d of the tube (separation of cathode and screen) is such that the electrons arrive in the same phase of the helix at the screen as that in which they left the photocathode.[18]

Both the above image tubes yield images of unity magnification. No such limitation exists if *short* electrostatic or magnetic lenses are employed, as shown in c and d of Fig. 2·14. In the first of these the lens consists of the accelerating field between two equidiameter cylinders, maintained at cathode and anode potential, respectively. The cathode and anode are flat disks terminating the two cylinders. Figure 2·15a

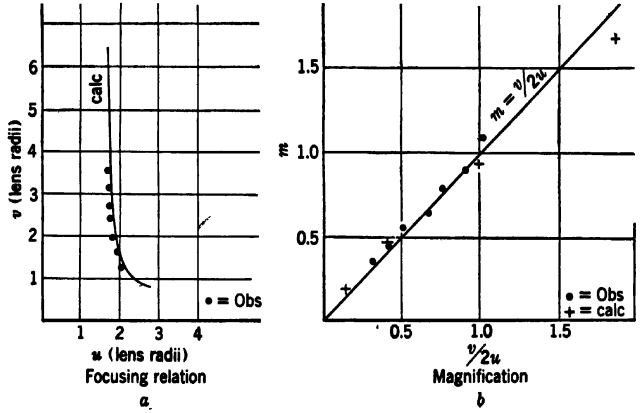


FIG. 2·15. Focusing Properties of the Electrostatic Image Tube Shown in Fig. 2·14c. (Courtesy *J. Optical Soc. Am.*, reference 10.)

shows the values of the distance v between the screen and the junction between the two cylinders (the *lens*) as function of the distance u between the cathode and the same junction. The magnification is given in each case by $v/(2u)$, as is shown in Fig. 2·15b.

The image formed by a tube of the type shown in Fig. 2·14c has a large amount of curvature of field, astigmatism, and distortion; hence it

[18] Thus, for focusing, the time of travel t from the cathode to the screen must be given by $vt = 1$. Since, at the same time, from the equations of motion of a uniformly accelerated body,

$$d = \frac{1}{2}\frac{e}{m}\frac{V}{d}t^2$$

the distance d must be related to the screen potential V and the magnetic field H by

$$d = \left(\frac{\pi c}{H}\right)\left(\frac{2mV}{e}\right)^{\frac{1}{2}} = \frac{11.05V^{\frac{1}{2}}}{H} \text{ cm} \qquad [2\text{·}6]$$

V being measured in volts and H in gauss.

is quite unsatisfactory in its outer portions. These defects are remedied by giving the cathode a radius of curvature approximately equal to u, the concave side being turned toward the screen. Furthermore, it is customary to replace the cathode cylinder by a series of rings connected to a lead passing through the tube envelope. By applying a small voltage V_1 with respect to the cathode to this lead the focusing of the tube can be varied after its construction has been completed. In this manner the electrostatic image tube shown in Fig. 2·16 is obtained.

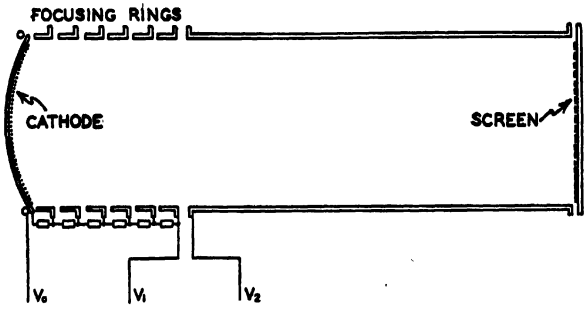


FIG. 2·16. Electrostatic Image Tube with Curved Cathode and Variable Focus. (Courtesy *J. Optical Soc. Am.*, reference 10.)

The magnetic image tube in Fig. 2·14*d* utilizes in effect a combination of electric and magnetic focusing fields. The magnification can be varied by shifting the focusing coil and altering the current therein, so as to maintain the sharpness of the image on the screen. Like the simple electrostatic tube, this magnetic tube possesses image defects, of which rotational distortion is perhaps the most serious. These defects can be reduced by employing special coils (Fig. 2·17) having the property that, in the neighborhood of the cathode, the magnetic field lines coincide with the paths that would have been traversed by electrons leaving the cathode with zero velocity in the absence of the magnetic field.

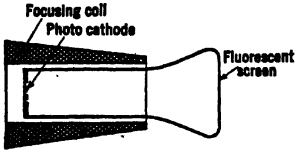


FIG. 2·17. Magnetic Image Tube with Focusing Coil Minimizing Aberrations. (Courtesy *Proc. Inst. Radio Engrs.*, reference 14.)

The image tubes shown in Fig. 2·14 have the common property of employing semitransparent cathodes and transparent screens. On both ends of the tube the efficiency is decreased somewhat by the absorption of light, in the cathode material and in the screen material, respectively,

and by the incomplete utilization of the incident light and electrons. The employment of superposed cylindrical electric and magnetic fields[19] (Fig. 2·18) makes this unnecessary. The optical image can be projected on the opaque photocathode, and the opaque fluorescent screen may be observed from the same side. The electrons now spiral about semicircles representing the magnetic field lines. Unfortunately, the image

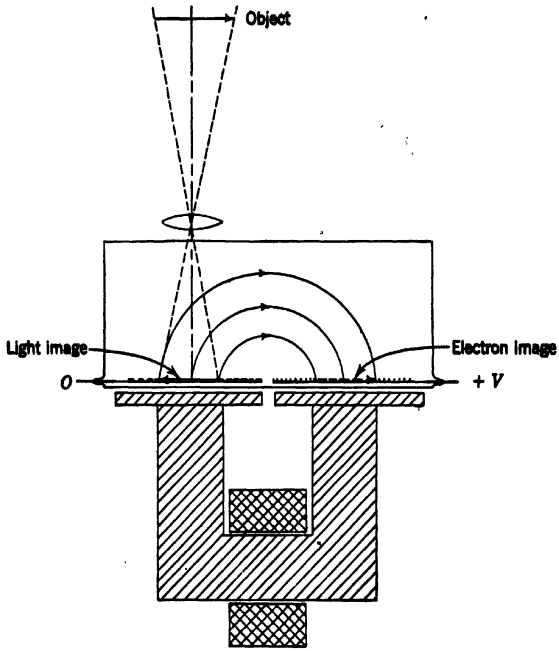


FIG. 2·18. Image Tube Employing Superposed Cylindrical Fields.

quality is not equal to that obtained with axially symmetric tubes, being marred, in particular, by the shear distortion visible in Fig. 2·19.

A characteristic application of an image tube of the type represented in Fig. 2·16 is shown in Fig. 2·20. Here an infrared image (or, for that matter, an ultraviolet image) of the specimen is focused by a light microscope on the photocathode of the image tube, the final image being viewed directly on the screen. Thus the image tube makes possible the direct observation of specimens illuminated with or emitting

[19] See Rose, reference 12

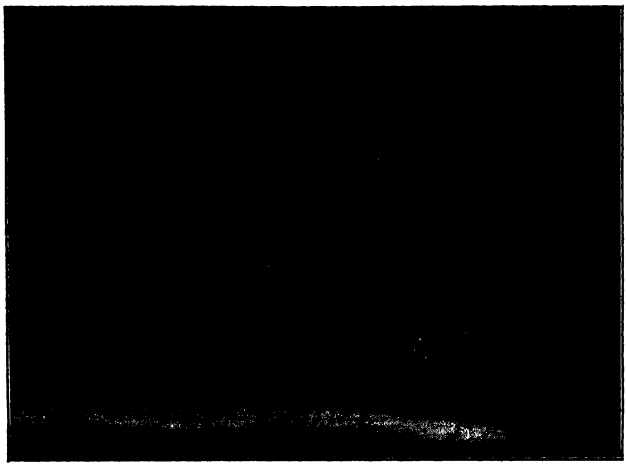


FIG. 2·19. Image of Resolution Pattern Produced by Image Tube Shown in FIG. 2·18. (Courtesy *Proc. Inst. Radio Engrs.*, reference 12.)



FIG. 2·20. Infrared Microscope Utilising Image Tube. (Courtesy *J. Optical Soc. Am.*, reference 10.)

invisible radiation. Figure 2·21 shows a micrograph of a daphnia, a small water creature, obtained in this manner, the specimen being illuminated with infrared radiation.



FIG. 2·21. Image of a Daphnia Obtained with Infrared Light. Magnification 20. (Courtesy *J. Optical Soc. Am.*, reference 10.)

Various special types of television pickup tubes have also made use of the image-tube principle. Thus Farnsworth's image dissector[20] (Fig. 2·22) consists essentially of an image tube, such as that shown in Fig. 2·14*b*, combined with a transverse deflecting field. The electron image as a whole scans a small aperture in the image plane; the electrons passing through the aperture constitute the picture signal and modulate, after amplification, the intensity of the scanning beam in the viewing tube. Furthermore, image tubes of the types shown in Figs. 2·16 and 2·17 have been combined with the more sensitive iconoscope

[20] See Farnsworth, reference 13.

so as to form the image iconoscope[21] or superemitron[22] tube, giving satisfactory performance at even lower light levels (Fig. 2·23). Here an electron image, rather than a light image, of the scene is projected on the sensitive mosaic of the iconoscope, storing in it charge in proportion to the brightness of the image which once in each picture period is released by the scanning beam.

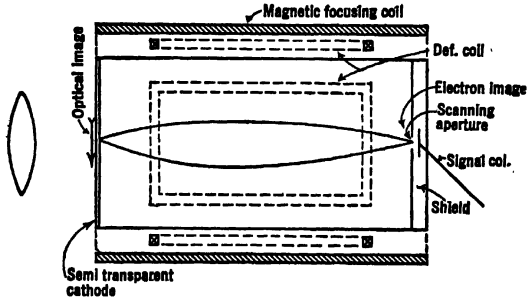


FIG. 2·22. Farnsworth Image Dissector (Early Type). (Zworykin and Morton, *Television*, John Wiley and Sons, New York, 1940.)

The repeated application of the image-tube principle offers the possibility of greatly enhancing the brightness of the final image obtained. The secondary-emission image multiplier represents one method of

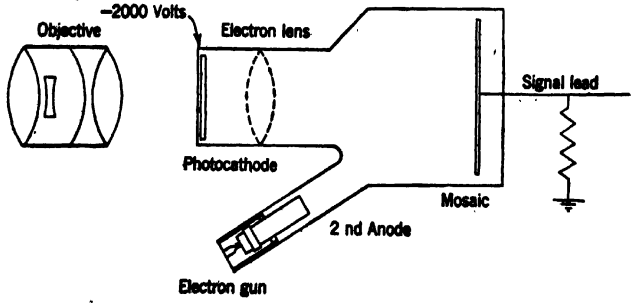


FIG. 2·23. The Image Iconoscope (Schematic Diagram). (Courtesy *Proc. Inst. Radio Engrs.*, reference 14.)

attaining this goal. Two realizations of this device are shown in Fig. 2·24. In the first of these (Fig. 2·24*a*) the image of the cathode *C* formed by the usual electrostatic image-tube configuration is projected

[21] See Iams, Morton, and Zworykin, reference 14.
[22] See "Superemitron Camera," reference 15.

on the upper half of the secondary-emission activated electrode T instead of on a fluorescent screen. This acts as cathode for a second

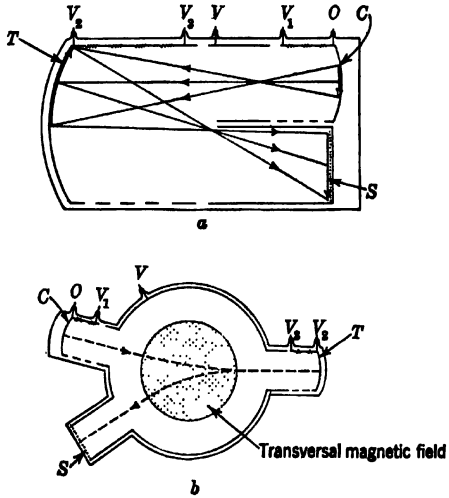


FIG. 2·24. Two Secondary-Emission Image Multipliers.

image tube with twice the diameter of the original one. In the second realization (Fig. 2·24*b*) a uniform magnetic field deflects the electrons coming from the cathode as well as the secondary electrons from the target T, so that the latter are directed into the neck containing the screen S rather than back toward the cathode. The images obtained with both arrangements suffer to some extent from distortion produced by the deviations from axial symmetry. A more basic difficulty of all secondary-emission image multipliers rests in the larger initial velocities of the secondary electrons as compared with those of thermionic or photoelectrons. These initial velocities give rise to large chromatic aberrations, rendering the final image unsharp. To some extent the chromatic aberrations can be compensated by the insertion of an electron mirror beyond the secondary-emission target, as shown in Fig. 2·25.

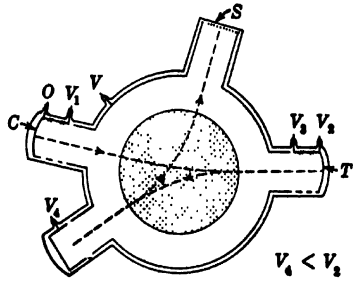


FIG. 2·25. Correction of Chromatic Aberration of an Image Multiplier by an Electron Mirror.

There exists another way of attaining an increase in image brightness by repeated application of the image-tube principle. This rests on the fact that, for a suitable choice of the photosensitive surface and the fluorescent screen, the photocurrent leaving such a surface placed next to the screen will exceed the high-velocity electron current bombarding the screen. If, under such circumstances, the electrons from the cathode are focused and accelerated onto a second similar screen, the image

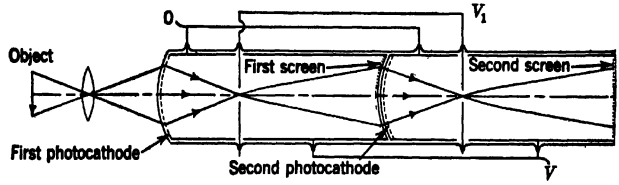


FIG. 2·26. Image Intensity Multiplier.

on the latter will be brighter than that on the original screen. Assuming reasonable values for the photosensitivity of the cathode and for the energy conversion efficiency of the luminescent screen, a gain in brightness may be expected with accelerating voltages of a few thousand volts.[23] In the practical realization the geometrical arrangement presents most difficulties, since the light from the fluorescent screens must be prevented from reaching the preceding cathodes. Figure 2·26 shows an arrangement of this type. Apertures at suitable potentials are employed to minimize the amount of illumination returning to the cathodes.

2·3. Deflection Systems. Electron guns and image-reproducing devices possess axial symmetry almost without exception. However, electron optics has acquired great importance also in fields where this is not the case. Thus deflecting systems, performing the function of variable electron prisms, are the basis of the success of electronic television. They find a similar application in oscillograph tubes.

The two basic types of deflecting fields are the uniform electrostatic field (Fig. 2·27) and the uniform magnetic field (Fig. 2·28). The former is parallel to the desired direction of deflection of the beam, the latter normal to this as well as to the beam itself.

In the uniform electrostatic deflecting field the paths of the electrons

[23] Let the sensitivity of the photocathode be $k = 20 \cdot 10^{-6}$ ampere per lumen for the emission of the fluorescent screen and let the screen have a luminous intensity of 2 candle powers per watt of incident electronic energy. The screen will then emit 2π lumens per watt in a forward direction. Then, if the accelerating voltage is V, the condition that the intensity of the second image be greater than that of the first is

$$2\pi kV > 1 \text{ or } V > 8000 \text{ volts}$$

are parabolas, corresponding to the trajectories of bodies which are thrown horizontally in a gravitational field. The ultimate angle of deflection is proportional to the length of the field and to the ratio of the field strength to the accelerating voltage of the tube.[24]

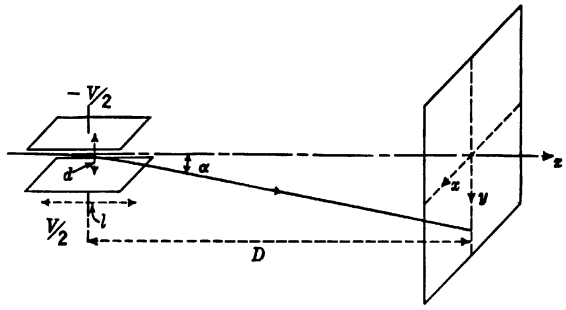


FIG. 2·27. Uniform Electrostatic Deflecting Field.

In the uniform magnetic field, on the other hand, the electrons describe circles, whose radius is proportional to the ratio of the electron velocity to the magnetic field. The total angle of deflection, being given by the ratio of the length of the field to this radius, is, accordingly, proportional

[24] Let V be the accelerating voltage of the tube, V_d the voltage applied between the deflecting plates, l the length of these plates, and d their separation. The initial velocity, parallel to the plates, will then be $v_z = (2eV/m)^{1/2}$. It will be maintained in the uniform field, so that the displacement in the z-direction is given, throughout, by

$$z = \left(\frac{2eV}{m}\right)^{1/2} t$$

t being the time measured from the instance of entry into the field (at $z = 0$, $y = 0$). In the y-direction, parallel to the deflecting field, the electrons experience, between the plates, a uniform acceleration $(e/m)V_d/d$. Hence the displacement in this direction is

$$y = \frac{eV_d t^2}{2md}$$

The total angle of deflection α is given by the ratio of the two components of velocity at $z = l$:

$$\tan \alpha = \frac{v_y}{v_z} = \frac{lV_d}{2dV} \qquad [2\cdot7]$$

expressing the time t in terms of $z = l$. The displacement at the screen, a distance D from the center of the deflecting field, becomes accordingly:

$$y_s = \frac{V_d l^2}{4dV} + \left(D - \frac{l}{2}\right)\frac{V_d l}{2dv} = \frac{l \cdot D \cdot V_d}{2dV} \qquad [2\cdot8]$$

to the length of the field, the field strength, and inversely proportional to the accelerating voltage of the electrons.[25] As for the uniform electrostatic field, the apparent center of deflection lies at the center of the actual deflecting field.

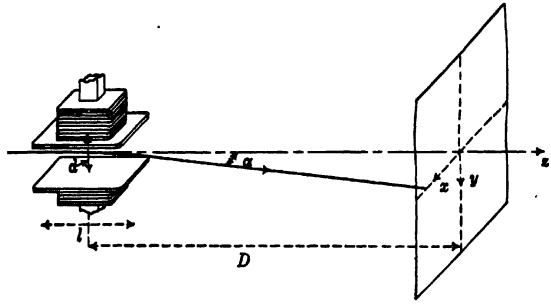


FIG. 2·28. Uniform Magnetic Deflecting Field.

In all actual cases the deflecting fields are nonuniform, since they must pass continuously over into a condition of zero field at great distance from the plates or coils. The resulting *fringe fields* cause deviations from the formulas given, resulting both in nonlinearity of the dependence of the displacement on the deflection voltage or current and in certain focusing effects of the beam. These focusing effects distort the spot as the beam is deflected, the portion of the beam nearer to the axis being subject to a slightly different field from the portion at greater distance

[25] The radius R of the path is determined by the condition that the acceleration normal to the path multiplied by the mass of the electron be equal to the force exerted by the magnetic field in the same direction:

$$\frac{mv^2}{R} = \frac{eHv}{c}$$

so that

$$R = \frac{mcv}{eH} = \left(\frac{2mc^2V}{eH^2}\right)^{1/2} = \frac{3.37V^{1/2}}{H}\text{ cm} \qquad [2\cdot9]$$

if V is measured in volts and H in gauss. The angle of deflection α is given, from the geometry of Fig. 2·28, by

$$\sin\alpha = \frac{l}{R} = \frac{0.297H}{V^{1/2}}l \qquad [2\cdot10]$$

and the displacement at the screen by

$$x_s = \left(D - \frac{l}{2}\right)\tan\alpha + l\tan\frac{\alpha}{2} \cong \frac{0.297lDH}{V^{1/2}}\text{ cm} \qquad [2\cdot11]$$

provided that $\alpha \ll 1$.

from the axis. These undesirable effects are largely overcome by shaping the deflecting elements appropriately.

In electrostatic deflection the deflecting elements are plates mounted on the inside of the tube, whose mean potential, preferably, is at all times equal to the anode potential. In order to achieve maximum deflection for a given applied voltage the plates are frequently flared so that they are just grazed by the edge of the beam at the maximum deflection (Fig. 2·29). If, as is usually the case, the beam is to be deflected in two mutually perpendicular directions, the pairs of plates are normally arranged in sequence, as in Fig. 2·29. If they are placed at the same point along the beam, a voltage applied to one pair will induce charges on the other, resulting in nonlinearity of deflection.

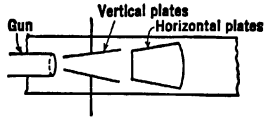


FIG. 2·29. Electrostatic Deflecting Plates Arranged in Sequence.

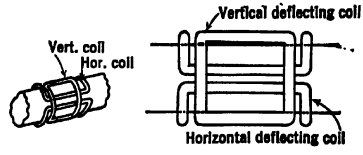


FIG. 2·30. Magnetic Deflection System with Air-Core Coils.

(Zworykin and Morton, *Television*, John Wiley and Sons, New York, 1940.)

With magnetic deflection this difficulty does not exist. The fields produced by the currents in two mutually perpendicular sets of deflecting coils are, to a first approximation, independent of each other. Thus the two coils are normally combined in a single *yoke* slipped over the outside of the cathode-ray tube. These yokes may be either iron-cored or air-cored. A system of air-cored coils yielding fields of a high degree of uniformity is illustrated in Fig. 2·30.

2·4. Beam Tubes and Multipliers. The application of electron optics to two-dimensional fields, that is, fields which do not vary appreciably in one direction, has also had a great influence in the development of modern amplifier tubes. By directing their attention to the paths of the individual electrons, abandoning the treatment of grids as two-dimensional surfaces of uniform potential, the designers of *beam tubes* have been able to improve the operating characteristics and to increase the efficiency and power output of amplifier tubes.

A typical modern amplifier tube or pentode is shown schematically, with an accompanying plot of the potential variation along a radius, in Fig. 2·31. Apart from the cathode, C, and the plate, P, there are three grids, G_1, G_2, and G_3. The first of these, the control grid, is common to all amplifier tubes. It is normally negatively biased; a small voltage variation applied to it causes a relatively large change in the current

which is capable of passing through it from the cathode to the plate. The plate, being connected through an external impedance to a source of positive potential, will drop in voltage as the current increases. The ratio of this drop in voltage to the increase in grid voltage causing it is the voltage amplification of the tube circuit. In modern tubes a screen grid G_2 is inserted between the plate and the control grid to reduce the control grid-plate capacity and thus to prevent feedback effects. This

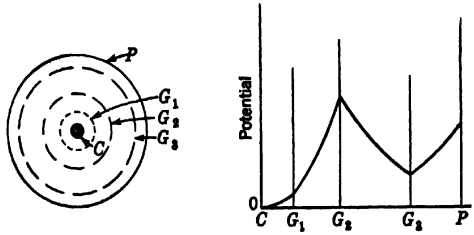


FIG. 2·31. Electrode Structure and Potential Variation in Pentode (Schematic).

is normally maintained at a fixed positive potential. The suppressor grid G_3, finally, is generally connected to the cathode and serves to return secondary electrons emitted by the plate to the latter, instead of letting them be collected by the screen grid. If the electron current is sufficiently high the space charge may form a *potential valley* between the screen grid and the plate, doing away with the need for a physical suppressor grid (Fig. 2·32).

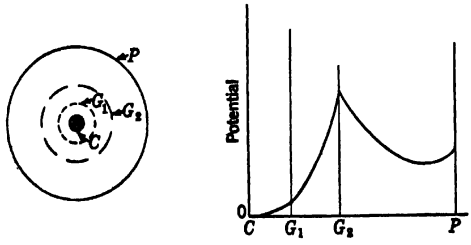


FIG. 2·32. Electrode Structure and Potential Variation in Tetrode with Potential Minimum Due to Space Charge Acting as Suppressor Grid (Schematic).

Unless special precautions are taken, a considerable portion of the electrons passing through the control grid will strike the screen grid. This will cause heating of the grid wires, which may prove a limiting factor in the power output of the tube. Furthermore, the proportion

of electrons absorbed by the screen grid will, in general, increase as the plate voltage decreases. This leads to a distorted output. Other undesirable consequences are reduced efficiency of the tube and enhanced random voltage fluctuations in the output, resulting from the distribution of the electrons between the plate and the screen grid. Thus one of the primary aims in the design of beam tubes is the reduction of screen-grid current by the proper arrangement of the grid wires relative to the cathode and, eventually, the fluting of the cathode itself (Fig. 2·33).[26]

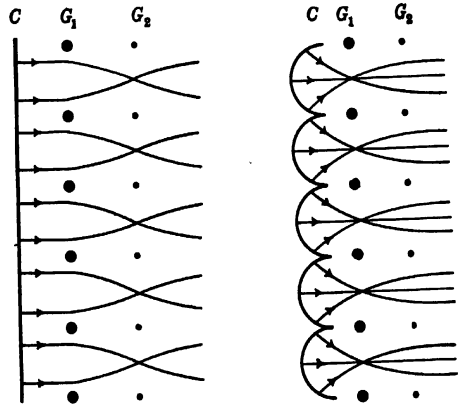


FIG. 2·33. Electrode Configurations for Minimizing Screen-Grid Current.

An excellent illustration of the application of electron-optical design principles is the beam power tube (RCA 6L6)[27] whose structure is shown, in cut-away section, in Fig. 2·34. The indirectly heated cathode is a flattened cylinder. The electrons leaving the flat side thereof are focused by the cylindrical lens fields between the aligned control-grid wires and screen-grid wires into flat pencils with a crossover midway between successive screen-grid wires. Electrons emitted by the narrow sides of the cathode are returned to it by the opposing field of the control-grid supports cooperating with the beam-confining plates (at cathode potential). Between the screen grid and the plate a potential minimum is formed by the space charge in the beam, suppressing secondary emission from the plate. This design not only prevents electrons from reaching the screen grid on their path toward the plate,

[26] See Thompson, reference 16, and Knoll, reference 17.
[27] See Schade, reference 18.

but also minimizes, as long as the potential minimum between screen grid and plate is positive, the number which are reflected by the minimum toward the screen grid. The electrons on their way toward the plate pass the screen-grid wires at a sufficient distance so that they are

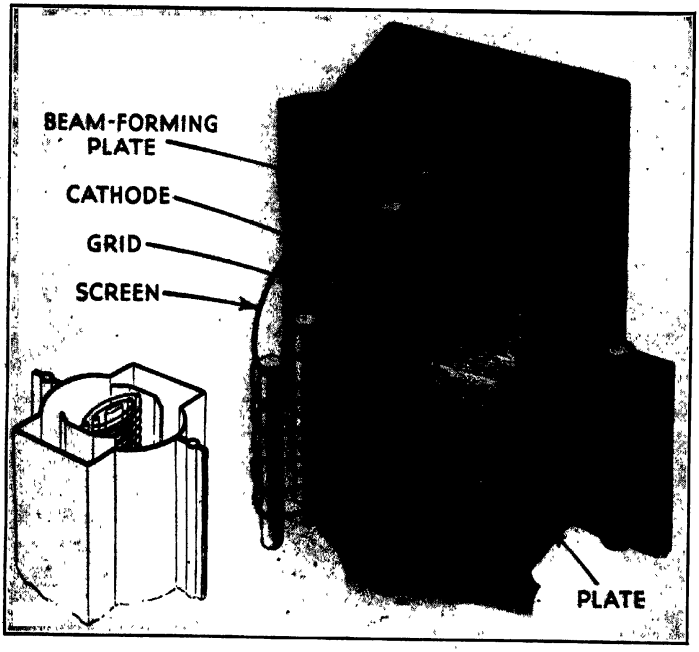


FIG. 2·34. Electrode Structure and Beam Formation in 6L6 Power Tube. (Courtesy *Proc Inst. Radio Engrs.*, reference 18.)

not given lateral deflections large enough to prevent them from passing on through the minimum to the plate. Figure 2·35 shows the potential distribution and the paths of the electrons in the structure for the limiting case in which the potential minimum between screen grid and plate is at zero potential, that is, in which it acts as a virtual cathode.

The confining of the beam by the beam-confining plates, the grid supports, and the shape of the cathode to a region of but slowly increasing horizontal width favors the formation of the secondary-emission suppressing potential minimum. The reduced screen-grid dissipation, on the other hand, prevents overheating of the control-grid wires and consequent thermionic emission of the grid.

A device whose development has been guided from the start by the methods of electron optics is the secondary-emission multiplier. Many surfaces — in particular oxidized metallic surfaces treated with alkali

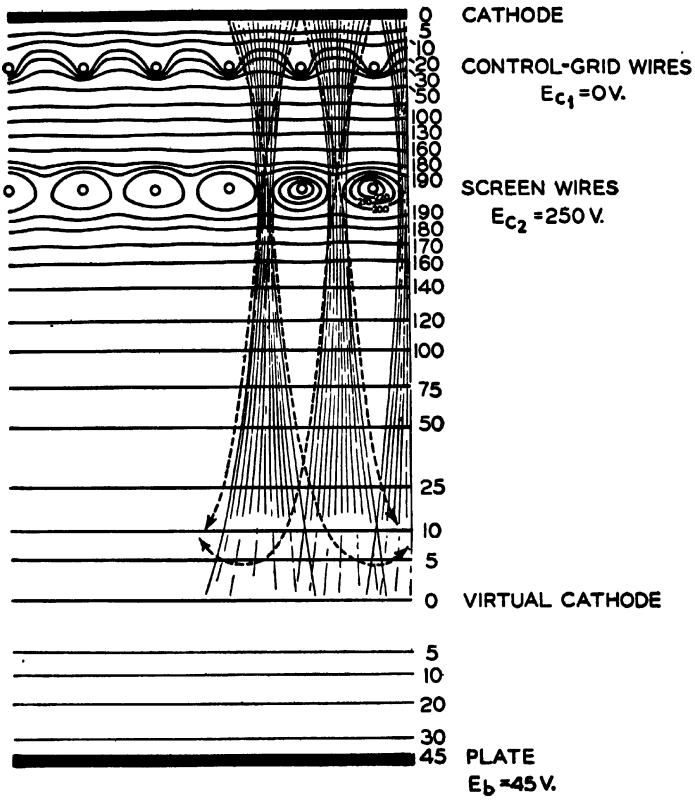


FIG. 2·35. Potential Variation and Electron Paths in 6L6. (Courtesy *Proc. Inst. Radio Engrs.*, reference 18.)

metals such as caesium or rubidium — have the property of emitting a number of low-velocity secondary electrons when a high-velocity (100- to 1000-volt) electron is incident on them. Thus the yield factor R, the ratio of the current of secondary electrons emitted by the surface to the incident primary electron current, for an appropriately oxidized and caesiated silver surface is plotted as function of the accelerating

voltage V of the incident electrons in Fig. 2·36. The high value of the yield factors may be utilized for amplifying weak electron currents, such as those from an illuminated photocathode, as shown schematically

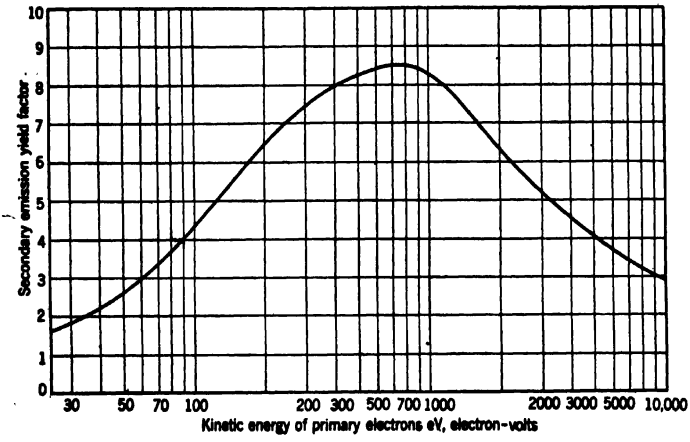


FIG. 2·36. Secondary-Emission Characteristic of a Typical Caesium-Activated Silver Surface.

in Fig. 2·37. Electrons from the cathode are accelerated to an activated target electrode and eject R times as many secondaries. These, in turn,

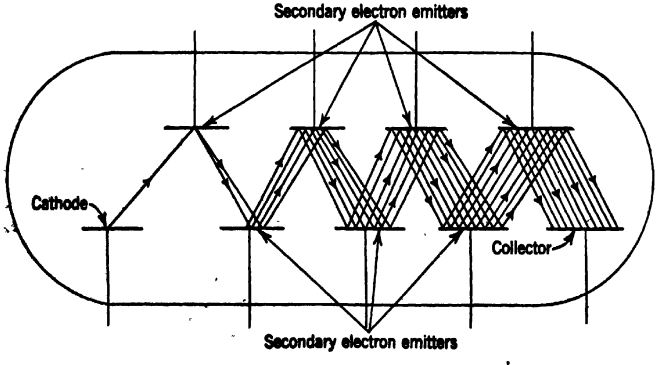


FIG. 2·37. Principle of the Secondary-Emission Multiplier.

are accelerated to the next electrode, ejecting from it R^2 as many electrons as originally left the cathode, provided that none of the electrons have been lost on the way. After n such current multiplications at n

successive targets, a current amplified by the factor R^n reaches the collector.

The primary conditions for the successful operation of a device of this type are (1) that the secondary electrons from one target reach the next without appreciable loss and (2) that these electrons, wherever incident, meet a retarding field which will draw the new secondaries away from the electrode and toward the next target.

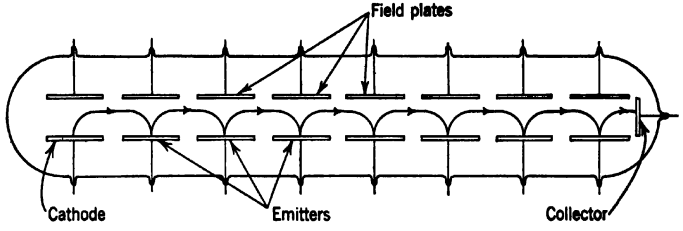


FIG. 2·38. Magnetic Secondary-Emission Multiplier (Schematic).

The first successful high-gain multiplier of this general character employed crossed magnetic and electric fields for this purpose (Fig. 2·38).[28] Electrons leaving any one of the activated plates, lying in a

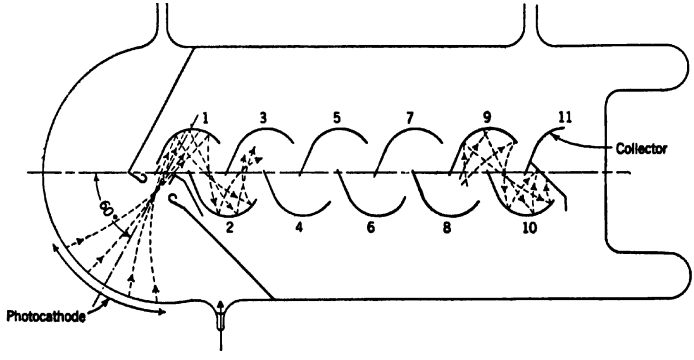


FIG. 2·39. Partition-Type Electron Multiplier. (Courtesy *Proc. Inst. Radio Engrs.*, reference 20.)

common plane, are accelerated toward the *field plate* opposite, this being connected electrically to the succeeding target electrode. They are prevented from reaching it by a magnetic field normal to the plane of the figure. This leads them, along an approximately cycloidal path, back

[28] See Zworykin, Morton, and Malter, reference 19.

toward the next emitter, giving rise to the ejection of new secondary electrons.

Although this multiplier proved very satisfactory in performance, it possessed the drawback of requiring a magnetic field, which is usually supplied by permanent magnets with soft-iron armatures. A careful study of the paths of electrons between curved electrodes at increasing

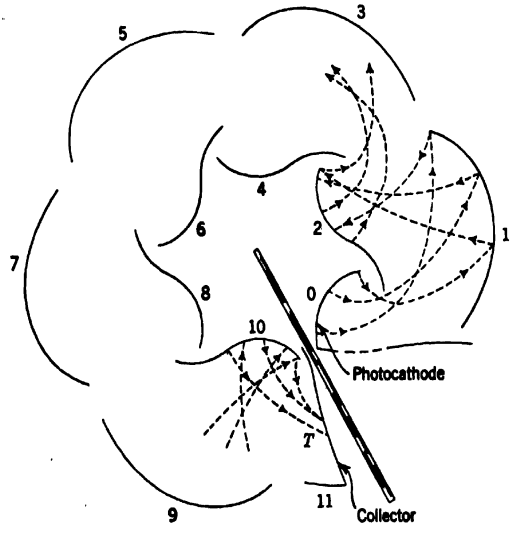


FIG. 2·40. Circular Electrostatic Electron Multiplier. (Courtesy *Proc. Inst. Radio Engrs.*, reference 20.)

positive potential in a staggered arrangement proved that the magnetic field could be omitted provided that the emitting plates were suitably shaped and placed. Figures 2·39 and 2·40 represent the structures of two purely electrostatic electron multipliers.[29] The first, especially designed for utilizing the photoemission from a relatively large photocathode, has a linear structure, like the magnetic multiplier previously discussed. In the second, on the other hand, the electrode structure has been wound up so as to fit into an envelope of the approximate size and shape of an ordinary radio tube (Fig. 2·41). Applying a difference of potential of 125 volts between two successive electrodes, or an overall voltage of 1250 volts between the photocathode and the collector, a total gain of 230,000 is obtained[30]; at the same time the dark current (the

[29] See Zworykin and Rajchman, reference 20.

[30] See Janes and Glover, reference 21.

output current due to thermionic and cold emission from the electrodes when no light falls on the photocathode) is normally of the order of one microampere. Larger gains and much smaller dark currents are obtained in special multipliers.

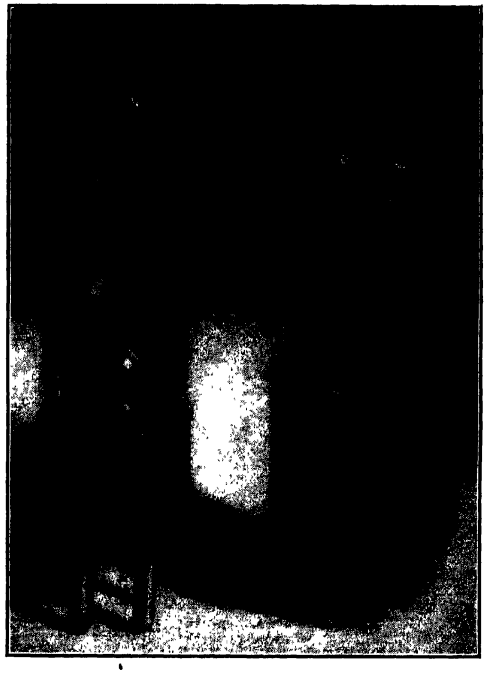


FIG. 2·41. External Appearance of Ordinary Phototube and Circular-Type Multiplier.

An interesting combination of the beam tube and the secondary-emission multiplier has been realized in the *orbital-beam multiplier* shown in Fig. 2·42.[31] Here the beams emitted by the cathode and grid structure are bent into circles by the electrostatic field between the inner (positive) and outer (negative) cylindrical electrode, striking the secondary-emission multiplier surface very close to the plate. Care has been taken in this design to make the time of travel from the cathode to the anode very nearly equal for all electrons. Thus the tube is well adapted for ultra-high-frequency work. The electron paths shown on

[31] See Wagner and Ferris, reference 22.

the right side of the central electrode illustrate the possibility of controlling the strength of the plate current by making the outer electrode more negative, causing an increasing fraction of the electrons to be collected by the central focusing electrode.

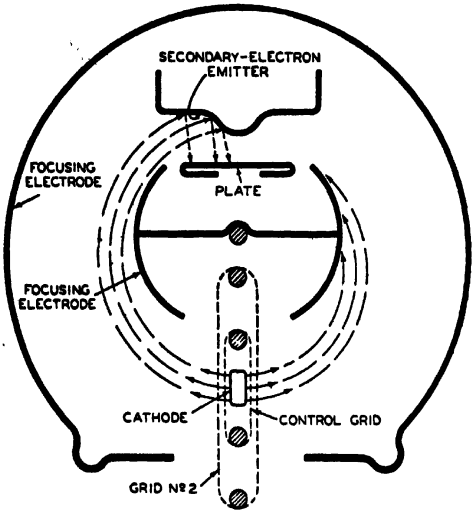


FIG. 2·42. Orbital Beam Multiplier. (Schematic Cross Section.) (Courtesy *Proc. Inst. Radio Engrs.*, reference 22.)

It is obvious that the secondary-emission multiplier phototube offers a means of attaining high gains with very compact equipment. Yet its value goes considerably beyond this. It introduces far less *noise*, or random fluctuations, into the photosignal transmitted by the electron current than a conventional resistance-coupled amplifier.[32] As a consequence it permits the satisfactory transmission of signals weaker by several orders of magnitude than those which can be transmitted with the aid of conventional means. These favorable noise characteristics have rendered the electron multiplier of great value in connection with television pickup tubes and in sound-film work. A further application will be noted in the discussion of the scanning electron microscope.[33]

2·5. Velocity and Mass Spectrographs. In light optics two different kinds of spectrographs, instruments which separate a given radiation

[32] For a discussion of the reduction in the noise output attainable by using a multiplier, see Appendix I.

[33] See section 3·4 and Appendix I.

into its components or its *spectrum*, are distinguished: grating spectrographs, depending on the change in the angle of diffraction with the wave length of the radiation, and prism spectrographs, depending on the variation of the refraction of a prism with wave length. Both also exist in the case of electron optics. In the first type the natural grating of crystalline matter[34] causes deflections of the incident beam proportional to the electron wave length. In the second, which is the subject of this section, electric and magnetic fields assume the role of prisms.

In light optics only one parameter distinguishes different radiations: the wave length. In particle optics (including electron optics) — extending the considerations to particles with charge and mass different

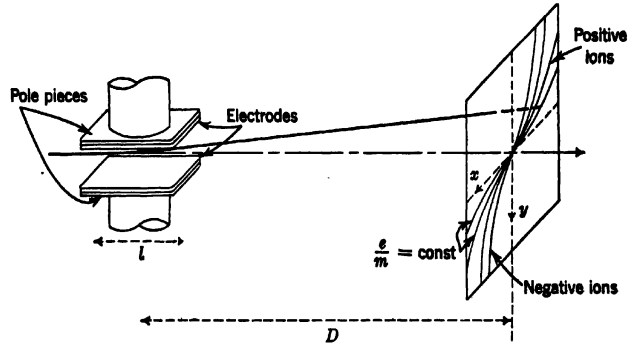


FIG. 2·43. Determination of e/m and Particle Velocity by the Parabola Method.

from the electron — there are three such parameters: the mass, the charge, and the velocity v, or, instead, the voltage drop V through which the particle has been accelerated. Two of these, the mass m and the charge e, cannot be separately determined by the methods of particle optics, since the motion of charged particles in electric and magnetic fields depends only on the ratio e/m. On the other hand, the actions of magnetic and electric fields on particles depend on different combinations of r and V with e/m, so that both the former quantities can be determined by a simultaneous application of both fields. Since there is good reason to believe that the quantity of charge of any particle must be an integer multiple of the electronic charge and since, furthermore, the integer factor is normally small, a measurement of e/m is usually regarded as a measurement of the mass m of the particle. An instrument which determines the values of e/m for particles composing a given beam is thus known as a *mass spectrograph*.

[34] See sections 1·1 and 19·9.

If a nonhomogeneous beam of charged particles is passed through a deflecting field consisting of superposed parallel electric and magnetic fields (Fig. 2·43), the particles will experience deflections in two mutually perpendicular directions. One, because of the electric field, will be inversely proportional to the voltage through which the particles have been accelerated; the other inversely proportional to the velocity of the particles, or the square root of the accelerating voltage. Thus particles with the same value of e/m, but having different velocities, will strike the screen, placed at some distance from the deflecting field, along parabolas. Particles with the same velocity, but with different values of e/m, on the other hand, fall on straight lines through the origin.[35]

More often than not this *parabola method* of J. J. Thomson[36] gives more information than is required — suffering, in compensation, from relatively low sensitivity. If, as is normally the case, the velocity distribution of a beam of identical particles *or* the mass spectrum[37] of an

[35] The dependence on e/m and v or V of the action of an electric and magnetic field on a particle is best illustrated by the expressions for the radius of curvature of a particle trajectory in a field normal to it:

$$R_e = \frac{mv^2}{eE} = \frac{2V}{E} \qquad [2\cdot12]$$

$$R_m = \frac{mcv}{eH} = \frac{c}{H}\left(\frac{2mV}{e}\right)^{\frac{1}{2}} \qquad [2\cdot13]$$

Here E is the strength of the electric field and H that of the magnetic field, and R_e and R_m are the radii of curvature in these two fields, respectively. Thus if two short, homogeneous, parallel electric and magnetic fields of length l are superposed and a beam is passed through them, the point of intersection of any particle with a plane a distance D from the compound field ($D \gg l$) will be given by

$$y = \frac{l \cdot D}{R_e} \qquad x = \frac{l \cdot D}{R_m} \qquad [2\cdot14]$$

y and x being rectangular coordinates with the point of intersection of the undeflected beam with the plane as origin.

Substituting from Eqs. 2·12 and 2·13 and solving for e/m, V, and v,

$$\frac{e}{m} = \frac{x^2}{y}\frac{c^2E}{lDH^2} \qquad V = \frac{lDE}{2y}$$

$$v = \frac{x}{y}\frac{cE}{H} \qquad [2\cdot15]$$

Particles of equal ratio of charge to mass thus come to lie on parabolas; those with equal velocity, on straight lines through the origin. These relations, with the constant coefficients suitably modified, hold also if the deflecting fields are not homogeneous.

[36] See reference 23.

[37] I.e., the relative representation of particles with different values of m/e in the beam.

inhomogeneous beam is sought, the velocity or mass spectrographs described below, utilizing various focusing principles, are advantageous.

The case is simple if only the velocity distribution of a beam of identical particles, for example, electrons, is to be investigated. The most generally useful method makes use of a uniform magnetic field. This focuses electrons of identical velocity leaving a fine slit in a slightly divergent pencil at a distance equal to twice the radius of curvature in the magnetic field (Eqs. 2·13 or 2·9) in a plane through the slit and normal to the direction of incidence (Fig. 2·44). Particle rays which deviate in angle from the principal ray strike the recording plane, which may be occupied by a photographic plate, at a distance slightly closer to the entrance slit, the deviation being of the second order in the angle of divergence.

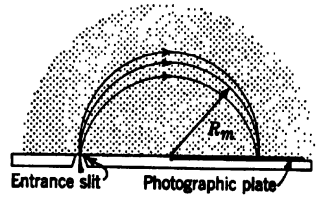


FIG. 2·44. Magnetic Velocity Spectrograph.

This type of velocity spectrograph has found wide application. It is particularly suitable for the determination of the velocity spectra of β-rays from radioactive substances, in which case the source itself, in the form of a narrow tube containing the radioactive material, may form the entrance slit.[38] In other cases, for example, the analysis of the photoelectrons ejected by a beam of x-rays from an extended target by Robinson,[39] the paths of the particles between the source and the slit are curved, so that different parts of the source contribute to the *spectrum lines* corresponding to different particle velocities. This may be avoided by replacing the photographic plate by a slit in fixed position, behind which a collector connected to an electrometer is placed. The electron distribution over the velocity spectrum is obtained by plotting the electrometer reading against the magnetic field strength. In Robinson's apparatus the magnetic field was produced by a pair of Helmholtz coils, that is, two similar coils with an equal number of turns with a radius equal to their separation.[40]

A radial electrostatic field formed between concentric cylindrical electrodes (Fig. 2·45) may be used instead of a magnetic field for obtain-

[38] See Ellis, reference 24.

[39] See reference 25.

[40] If the radius of the coils is a, the very nearly uniform field at the center between the coils is given by (see Starling, reference 26, p. 57):

$$H = 0.899\frac{ni}{a} \text{ gauss} \qquad [2\text{·}16]$$

where i is the current through the coils in amperes.

ing a velocity spectrum. In this case the radius of curvature of the paths is given by Eq. 2·12. Hughes and Rojansky[41] have shown that such a field brings electrons to focus after a deflection through an angle $\pi/2^{\frac{1}{2}} = 127° 17'$. Although this method is less convenient than the magnetic method whenever the latter can be applied, it is capable of giving approximately the same resolution.[42]

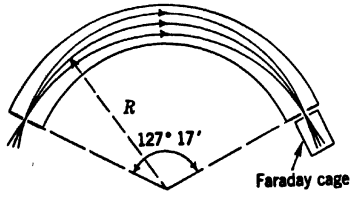


FIG. 2·45. Electrostatic Velocity Spectrograph.

Velocity spectrometers of quite different type, sacrificing resolution to sensitivity, make use of the variation in the focal length of a magnetic

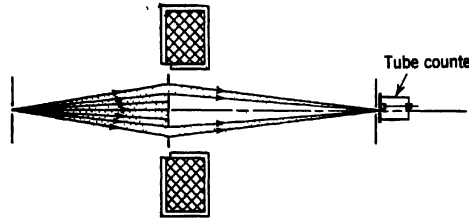


FIG. 2·46. Magnetic Lens Spectrograph. (Klemperer.)

lens with applied voltage (Fig. 2·46).[43] This focal length is, roughly, proportional to the kinetic energy eV of the particles and inversely proportional to the square of the magnetic field strength within the lens.[44]

[41] See reference 27.

[42] If Δv is the least velocity difference which will give rise to two distinct spectrum lines and d_s is the width of the entrance slit, the optimum resolution is given in either case by

$$\frac{\Delta v}{v} \simeq \frac{d_s}{2R} \qquad [2·17]$$

where R is the radius of curvature of the path.

[43] See Klemperer, reference 28.

[44] The focal length of a short magnetic lens, for arbitrarily high particle velocities, is given by

$$f = \frac{\frac{8mc^2}{e}\left(V + \frac{eV^2}{2mc^2}\right)}{\int H_z^2\,dz} = \frac{4(HR_m)^2}{\int H_z^2\,dz} \qquad [2·18]$$

Here eV is the kinetic energy of the particles, H_z the magnetic field along the axis of the lens, and H the magnetic field required to give the path of the electron a radius of curvature R_m. HR_m is a frequently employed measure of the momentum of high-speed particles, being obtained directly from cloud-chamber observations.

Thus, if the magnetic field strength is varied, particles of different kinetic energy come to focus at the same point on the axis of the lens. To measure the magnitude of the particle currents corresponding to different velocities, a Geiger-Müller counter or some other sensitive current-measuring device is placed at this point. The central area of the lens is blocked, leaving an annular region clear. In Klemperer's apparatus almost one-thousandth of the total radiation of given velocity emitted by the source could be utilized. More recently, Coslett[45] has adapted it for the analysis of electrons in the range of 15 million electron volts.

Klemperer[43] has also built another lens spectrograph, suitable for attaining higher resolutions. As shown in Fig. 2·47, this instrument is

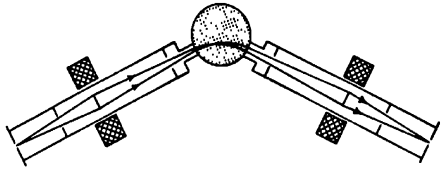


FIG. 2·47. Compound Magnetic Velocity Spectrograph. (Klemperer.)

almost the perfect analogue of the usual prism light spectrometer, the prism being replaced by a uniform magnetic field normal to the plane of the figure.

The mass spectrograph has played — and is playing — possibly an even greater role in the history of physics than the velocity spectrograph. Thus it has furnished the experimental basis for the concept that the atomic nuclei themselves are made up of more elementary building stones. It is clear from Eq. 2·13 that any magnetic velocity spectrograph may serve as mass spectrograph, provided that all the ions in the beam are accelerated through the same drop of potential, that is, have the same value of V. Thus Dempster[46] has developed the magnetic spectrograph with 180-degree focusing (Fig. 2·44) into a precision instrument for that purpose. The source of ions is here a suitable salt deposit on an electrode, which is bombarded by electrons from a near-by filament.

This earlier mass spectrograph of Dempster was capable of focusing particles with different initial directions of motion, but equal velocity and equal m/e. Aston,[47] on the other hand, constructed an instrument

[45] See reference 29.
[46] See reference 30.
[47] See reference 31.

which, while requiring a substantially parallel initial beam defined by two narrow, well-separated apertures, brings particles of slightly varying velocities, but having the same value of m/e, to focus at the same point. Aston's instrument also differs from Dempster's in using a photographic plate in place of a collecting electrode for recording the spectrum.

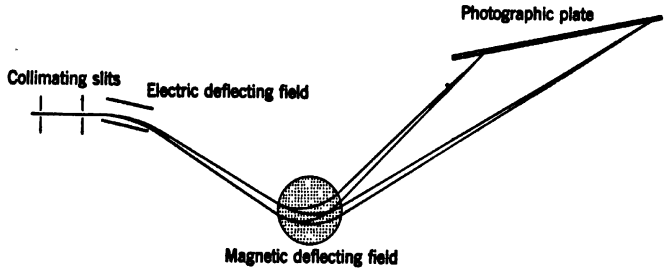


FIG. 2·48. Principle of Aston's Mass Spectrograph.

The basic principle of Aston's instrument, shown schematically in Fig. 2·48, consists in first deflecting the ions in an electrostatic field and then overcompensating the deflection, as far as particles of the same m/e, but different velocity, are concerned, by a magnetic field. There exists then a point for each m/e at which particles of different velocity intersect. These points do not coincide, since only the magnetic field discriminates between particles of different m/e and identical V. With a suitable choice of the position and strength of the electric and magnetic fields the focusing points come to lie on a straight line, defining the proper position of the photographic plate. Furthermore, the position of the " spectrum lines " on the plate varies linearly with m/e.

The problem of the mass spectrograph cannot be regarded as completely solved, however, until particles with the same m/e, but differing in both velocity and initial direction, are focused in a single line. The electric and magnetic fields must act at the same time as the prism and as an achromatic cylindrical lens. That this is possible was demonstrated by Mattauch and Herzog.[48] Employing a radial electric field as " collimator " and a magnetic field of special shape both as prism and as " telescope lens " compensating the chromatic error of the collimator, they built a spectrograph (Fig. 2·49) which focuses for both velocity differences and angular divergence for all values of m/e. The resolution achieved with this instrument is 1:6000. An even greater resolution has been reported by Bainbridge and Jordan[49] for an instrument com-

[48] See Mattauch and Herzog, reference 32, and Mattauch, reference 33.
[49] See reference 34.

bining a 127-degree radial electric field with a 60-degree magnetic field. This fulfils the condition of *double focusing* for one value of m/e and that of velocity focusing alone throughout. Whereas the scale of Mattauch and Herzog's instrument is proportional to $(m/e)^{\frac{1}{2}}$, that of Bainbridge and Jordan's instrument is linear. Other, different, field combinations realizing at least partial double focusing have found application in more recent apparatus of Dempster[50] and Aston.[51]

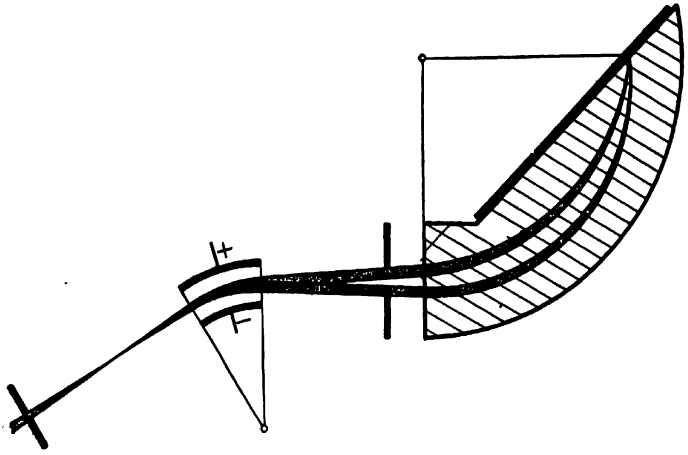


FIG. 2·49. Double-Focusing Mass Spectrograph. (Mattauch and Herzog, reference 32.)

2·6. Particle-Accelerating Devices. With the increasing interest in the constitution of the atomic nuclei, the demand for charged particles of very high energy, which could be used for disintegration experiments, became progressively more insistent. It has been met primarily in two ways. On the one hand, new methods have been developed for the generation of very high voltages,[52] which permit the direct acceleration of charged particles by passage from one electrode to another electrode of opposite potential or to ground. On the other hand, alternating fields have been utilized to give particles the required accelerations without the employment of comparably high voltages. Both systems involve interesting problems in particle optics. In the first they are largely limited to holding the beam together in the long accelerating

[50] See reference 35.

[51] See reference 36. For more recent developments in mass spectroscopy, the reader is referred to Jordan *et al.*, reference 37.

[52] See Tuve, Hafstad, and Dahl, reference 38.

tube, the great length being required to prevent breakdown on the outside and cold emission on the inside. The focusing effect is here attained by passing the beam through a succession of electrostatic lenses formed between cylinders at increasing potentials (Fig. 2·50). The relative voltages of the several electrodes are partly controlled by the setting of corona points attached to the external shield of each electrode.

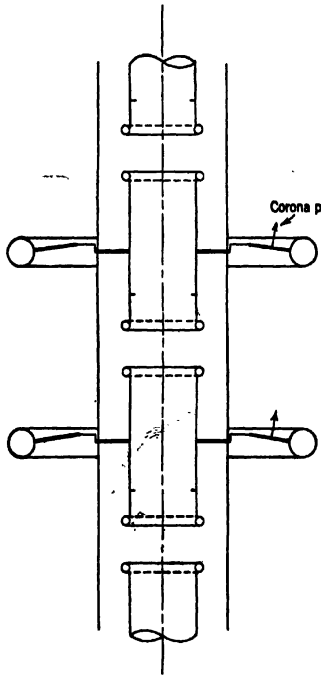


FIG. 2·50. Section of Accelerating Tube. (Tuve, Hafstad, and Dahl. Courtesy *Phys. Rev.*, reference 38.)

Tubes quite similar to these accelerating tubes can also be used for obtaining high-velocity particles without the use of very high voltages. Thus the linear accelerator of Sloan and Lawrence[53] consists essentially of an ion source followed by a sequence of cylindrical electrodes of increasing length and, finally, a target or a Faraday cage for collecting the ions. However, instead of a high voltage being applied between the source and the target, an alternating voltage is applied between the sets of odd- and even-numbered electrodes, the frequency of alternation and the electrode lengths being so chosen that for particles (ions) entering in the proper phase of the voltage wave the field encountered between any pair of electrodes is an accelerating field of maximum amplitude.[54] The length of the electrodes then becomes proportional to the square root of their order number and inversely proportional to the frequency and to the square root of the ratio of the mass of the particles to their charge. Since the frequency cannot be increased indefinitely,

[53] See reference 39.

[54] If v_n is the velocity of the particle in the nth electrode, of length L_n, ν is the frequency of alternation, and V_a is the amplitude of the alternating voltage, the condition for continuous acceleration becomes

$$\nu = \frac{v_n}{2L_n} \quad \text{or} \quad L_n = \frac{v_n}{2\nu} = \frac{1}{\nu}\left(\frac{enV_a}{2m}\right)^{\frac{1}{2}} \qquad [2\cdot19]$$

a moderate length of the accelerating tube is most readily attained by employing ions of great mass.

In Sloan and Lawrence's apparatus (Fig. 2·51) the particles were mercury ions formed in a hot-cathode arc discharge. These were drawn through the perforated negative electrode A and the focusing cylinder B into the accelerator system of 30 stages. To this an alternating potential with an amplitude of 42,000 volts and a frequency of about 10 megacycles per second was applied. The overall length of

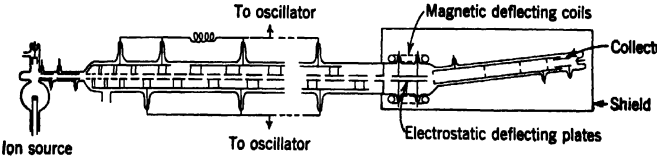


FIG. 2·51. Linear Accelerator for Heavy Ions. (Sloan and Lawrence. Courtesy *Phys. Rev.*, reference 39.)

the tube was 114 centimeters. At the high-velocity end of the accelerator an electrostatic field serves to direct the ions of the desired velocity into a Faraday cage. Electrons which may pass through or be generated in the system are removed by a magnetic deflecting field, which has no appreciable effect on the much heavier ions. With this apparatus ion currents of 10^{-7} ampere were obtained with a particle energy of 1.26 million electron volts. Slightly smaller currents of ions with an energy of 2.89 Mev were collected in a later instrument by Sloan and Coates.[55]

The linear accelerator has certain limitations: (1) It becomes excessively long and cumbersome for the light ions, which are most effective for nuclear disintegrations, and also for a large voltage amplification factor (that is, a large number of stages). (2) It utilizes only a small fraction of the ions emitted, namely, those entering the system at the peak of the voltage wave. (3) The focusing conditions for the beam become less favorable as the length, and hence the voltage amplification, are increased.[56] These difficulties are largely overcome in the *cyclotron* of Lawrence and his coworkers,[57] which, at the present time, is unquestionably the most effective source of high-speed particles for disintegration experiments.

In the cyclotron there are only two, D-shaped, electrodes, between

[55] See reference 40.

[56] See section 18·5.

[57] See, e.g., Lawrence and Livingston, reference 41; Livingston, reference 42; and Lawrence and Cooksey, reference 43.

which the alternating voltage is applied. A magnetic field bends the paths of the charged particles in the field-free region within the electrodes into semicircles, so that particles which have crossed the gap between the electrodes return to the latter in just one-half period (Fig. 2·52). Thus particles which are accelerated in the original transit through the gap continue to be accelerated to the same degree in every succeeding transit. Since, with increasing particle velocity, the radius of the semicircular paths increases proportionately, the path of any one particle takes on the appearance of a spiral. Ultimately, when the value of R_m (Eq. 2·13) approaches the radius of the *dees*, the particles pass through a slit into a strong electric deflecting field, which draws them out of the magnetic field toward any selected target. Since the frequency of rotation of a charged particle is given simply by the product of the magnetic field strength and the ratio of the charge to the mass of the particle,[58] it is merely necessary to set the frequency of the alternating potential on the two electrodes equal to this value. The particles will then be accelerated in the same phase at every successive transit between electrodes, regardless of their phase of entrance.

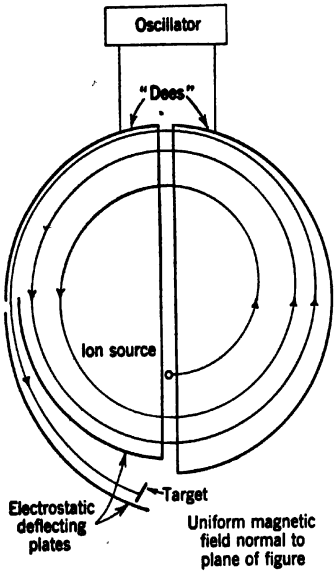


FIG. 2·52. Principle of the Cyclotron.

Figure 2·53 shows the external appearance of the 60-inch Berkeley cyclotron.[59] The large cylinders in front represent the tank circuit connected to, and supporting, the dees. It is connected to the oscillator (far to the right) by means of the two concentric lines formed by the large tubes sloping upwards to the right. The sheet-metal-covered box in the background contains the Lecher wires of the oscillator. The figure is leaning against the magnet, which, in part, is hidden by the

[58] Thus for protons (hydrogen nuclei of unit atomic mass) the frequency of rotation is

$$\nu = \frac{eH}{2\pi mc} = 1.53 \cdot 10^8 H \text{ sec}^{-1} \qquad [2\cdot20]$$

[59] See Lawrence *et al.*, reference 44.

tank circuit. The black cable at the left conducts the high voltage to the electrostatic deflecting plate, which pulls the accelerated particles out of the cyclotron field. The vacuum chamber containing the two dees, which is placed between the pole pieces of the magnet, is barely visible. Between the dees an alternating voltage of some 100,000 volts amplitude is applied. These are flat semicircular chambers providing



FIG. 2·53. The 60-Inch Berkeley Cyclotron. (Courtesy Dr. E. O. Lawrence.)

electric shielding for the ions traveling within them. In order to generate the ions the whole vacuum vessel is filled with light or heavy hydrogen or with helium at a low pressure, and hot filaments emitting electrons are placed above and below the gap. The electrons traveling, owing to the presence of the strong magnetic field, in a narrow column normal to the dees ionize the gas in the gap, the ions being pulled by the alternating field into the interior of the dees. As described in greater detail in section 18·5, both the lens properties of the electrostatic field at the gap and the curvature of the magnetic field lines near the edge of the pole pieces are utilized to prevent the loss of ions by their colliding with the top and bottom of the circulating chambers. Even in the first experiments with the instrument discussed above, beams of 10 microamperes of deuterons (heavy-hydrogen nuclei, twice as heavy as protons or ordinary, light-hydrogen, nuclei) with an energy of 16 million electron volts (Mev) were obtained.

A third accelerating device, suited primarily for furnishing electrons

of extremely high velocity, depends on an entirely different process. The induction accelerator or betatron, analyzed theoretically by Wideroe[60] as early as 1928, but developed into a working instrument only relatively recently by Kerst,[61] makes use of the familiar fact that a changing magnetic field is accompanied by an electric field, which, in the presence of a conductor, sets up within the latter a current tending to compensate the change in magnetic field. If free charges are present

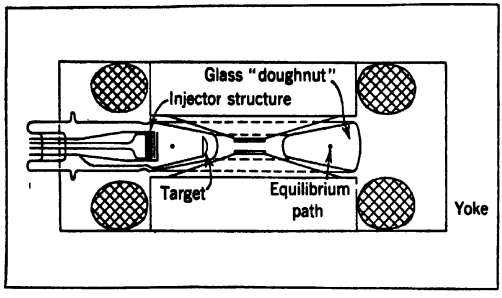


FIG. 2·54. The Magnetic Induction Accelerator. Schematic Diagram of the 2.3-Mev Instrument.

in place of the conductor, this electric field will accelerate them. Since, in an axially symmetric magnetic field, the electric field is normal both to the radius vector to the axis and to the axis of the magnetic field itself, it might be expected to cause the charged particle to spiral outward at a rapid rate. However, under certain conditions the particle will be accelerated continuously on the same circular orbit. They are:[62]

1. The magnetic field at the orbit must at all times be just half as large as the average magnetic field in the circular area bounded by the orbit.
2. The initial velocity of the particles must be such that the center of curvature of their path falls on the axis of the magnetic field.

If these requirements are satisfied, the particle will increase in velocity as long as the magnetic field increases and will decrease in velocity when it decreases, still remaining on the same orbit.

The general arrangement of Kerst's first induction accelerator is shown in Fig. 2·54. Electrons from the injector structure, compris-

[60] See reference 45.
[61] See references 46 and 47.
[62] See section 13·5.

ing a helical tungsten filament and focusing and accelerating electrodes, enter an internally silvered glass *doughnut* which is placed between the poles of an electromagnet. This is constructed so that the field at the position of the equilibrium orbit is just half the mean field in the circular area bounded by this orbit. Leaving the injector structure with a low velocity, corresponding to 100–1000 volts, when the field between the poles is nearly zero, the electrons spiral inward toward the equilibrium orbit, clearing the injector structure. Here they remain,

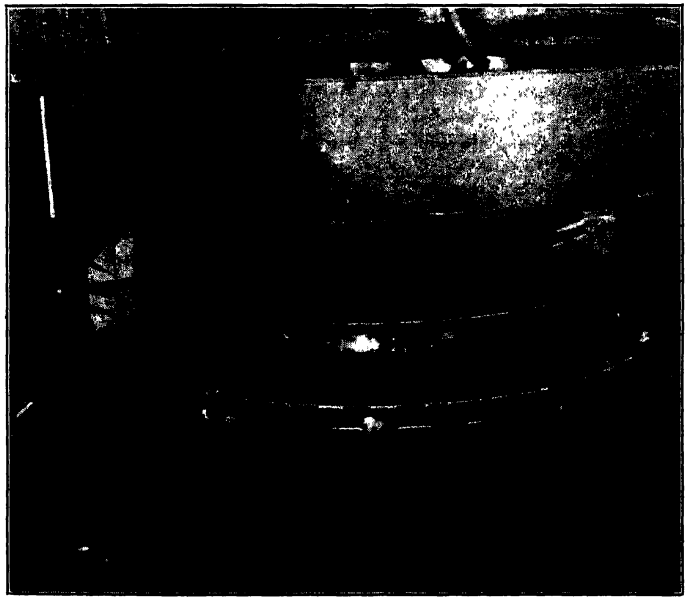


FIG. 2·55. The 20-Mev Betatron; Upper Portion of Magnet Raised to Show Vacuum Chamber. (Courtesy Dr. D. W. Kerst and *Rev. Sci. Instruments*, reference 47.)

gaining constantly in velocity, until the magnetic field approaches its maximum value (in about 1/2400 second, the operating frequency of the electromagnet being 600 cycles per second). Near this point the central core becomes magnetically saturated, causing the mean magnetic field within the orbit to increase less rapidly, so that the tangential force on the particle lags behind the centripetal force exerted by the magnetic field at the orbit. As a consequence the electrons spiral inward and strike a target, whose resulting x-ray emission is measured.

A current of about 0.02 microampere of 2.3 Mev electrons was obtained with this instrument.

In the new instrument[63] (Fig. 2·55) the energy of the electrons has been raised to 20 Mev. Here, instead of spiraling inward at the end of their course, the electrons are drawn out again so as to bombard the injector structure. This is accomplished with the aid of a strong current pulse through a coil arranged so as to enhance the magnetic flux through the center of the apparatus. Various improvements in design and technique of operation have made possible an increase in the output current commensurable with the increase in particle energy.

REFERENCES

1. V. K. Zworykin, "The iconoscope — a modern version of the electric eye," *Proc. Inst. Radio Engrs.*, Vol. 22, pp. 16–32, January 1934.
2. V. K. Zworykin, "Description of an experimental television system and the kinescope," *Proc. Inst. Radio Engrs.*, Vol. 21, pp. 1655–1673, December 1933.
3. V. K. Zworykin, "Television with cathode-ray tube for receiver," *Radio Eng.*, Vol. 9, No. 12, pp. 38–41, December 1929.
4. V. K. Zworykin and W. H. Painter, "Development of the projection kinescope," *Proc. Inst. Radio Engrs.*, Vol. 25, pp. 937–953, August 1937.
5. R. R. Law, "High-current gun for projection kinescope," *Proc. Inst. Radio Engrs.*, Vol. 25, pp. 954–976, August 1937.
6. D. W. Epstein, "Electron-optical system of two cylinders as applied to cathode-ray tubes," *Proc. Inst. Radio Engrs.*, Vol. 24, pp. 1095–1139, August 1936.
7. V. K. Zworykin, U. S. Patent 2,084,364, filed December 24, 1930.
8. H. Iams, "A fixed-focus electron gun for cathode-ray tubes," *Proc. Inst. Radio Engrs.*, Vol. 27, pp. 103–105, February 1939.
9. R. B. Janes and W. H. Hickok, "Recent improvements in the design and characteristics of the iconoscope," *Proc. Inst. Radio Engrs.*, Vol. 27, pp. 535–540, September 1939.
10. V. K. Zworykin and G. A. Morton, "Applied electron optics," *J. Optical Soc. Am.*, Vol. 26, pp. 181–189, April 1936.
11. G. Holst, J. H. de Boer, M. C. Teves, and C. F. Veenemans, "Transformation of light of long wavelength into light of short wavelength," *Physica*, Vol. 1, pp. 297–305, February 1934.
12. A. Rose, "Electron optics of cylindrical electric and magnetic fields," *Proc. Inst. Radio Engrs.*, Vol. 28, pp. 30–40, January 1940.
13. P. T. Farnsworth, "Television by electron image scanning," *J. Franklin Inst.*, Vol. 218, pp. 411–444, October 1934.
14. H. Iams, G. A. Morton, and V. K. Zworykin, "The image iconoscope," *Proc. Inst. Radio Engrs.*, Vol. 27, pp. 541–547, September 1939.
15. "Superemitron Camera," *Wireless World*, Vol. 41, pp. 497–498, November 1937.
16. H. C. Thompson, "Electron beams and their application in low-voltage devices," *Proc. Inst. Radio Engrs.*, Vol. 24, pp. 1276–1297, October 1936.
17. M. Knoll, "The amplifier and transmitter tube as electron-optical problem," *Z. tech. Physik*, Vol. 15, pp. 584–591, December 1934.

[63] See Kerst, reference 47.

18. O. H. SCHADE, " Beam power tubes," *Proc. Inst. Radio Engrs.*, Vol. 26, pp. 137–181, February 1938.
19. V. K. ZWORYKIN, G. A. MORTON, and L. MALTER, " The secondary-emission multiplier — a new electronic device," *Proc. Inst. Radio Engrs.*, Vol. 24, pp. 351–375, March 1936.
20. V. K. ZWORYKIN and J. A. RAJCHMAN, " The electrostatic electron multiplier," *Proc. Inst. Radio Engrs.*, Vol. 27, pp. 558–566, September 1939.
21. R. B. JANES and A. M. GLOVER, " Recent developments in phototubes," *RCA Review*, Vol. 6, pp. 43–54, July 1941.
22. H. M. WAGNER and W. R. FERRIS, " The orbital-beam secondary-emission multiplier for ultrahigh-frequency amplification," *Proc. Inst. Radio Engrs.*, Vol. 29, pp. 598–602, November 1941.
23. J. J. THOMSON, " Rays of positive electricity," *Phil. Mag.*, Vol. 21, pp. 225–249, 1911.
24. C. D. ELLIS, " Magnetic spectrum of β-rays excited by γ-rays," *Proc. Roy. Soc.*, London, Vol. A99, pp. 261–271, June 1921.
25. H. ROBINSON, " Secondary corpuscular rays produced by homogeneous x-rays," *Proc. Roy. Soc.*, London, Vol. A104, pp. 455–479, November 1923.
26. S. G. STARLING, *Electricity and Magnetism*, Longmans, Green and Co., London, 1926.
27. A. LL. HUGHES and V. ROJANSKY, " On the analysis of electron velocities by electrostatic means," *Phys. Rev.*, Vol. 34, pp. 284–290, July 1929.
28. O. KLEMPERER, " The use of electron lenses for β-rays," *Phil. Mag.*, Vol. 20, pp. 545–561, October 1935.
29. V. E. COSLETT, " Magnetic lens for β-rays of high energy," *J. Sci. Instruments*, Vol. 17, pp. 259–264, July 1940.
30. A. J. DEMPSTER, " A new method for positive-ray analysis," *Phys. Rev.*, Vol. 11, pp. 316–325, April 1918.
31. F. W. ASTON, " Positive-ray spectrograph," *Phil. Mag.*, Vol. 38, pp. 707–714, 1919.
32. J. MATTAUCH and R. HERZOG, " A new mass spectrograph," *Z. Physik*, Vol. 89, pp. 786–795, July 1934.
33. J. MATTAUCH, " A double-focusing mass spectrograph and the masses of N^{15} and O^{18}," *Phys. Rev.*, Vol. 50, pp. 617–623, October 1936.
34. K. T. BAINBRIDGE and E. B. JORDAN, " Mass spectrum analysis," *Phys. Rev.*, Vol. 50, pp. 282–296, August 1936.
35. A. J. DEMPSTER, " New methods in mass spectroscopy," *Proc. Am. Phil. Soc.*, Vol. 75, pp. 755–766, 1935.
36. F. W. ASTON, " A second-order focusing mass spectrograph and isotopic weights by the doublet method," *Proc. Roy. Soc.*, London, Vol. A163, pp. 391–404, December 1937.
37. E. B. JORDAN, L. B. YOUNG, N. D. COGGESHALL, J. A. HIPPLE, and D. RITTENBERG, Special Issue on Mass Spectroscopy, *J. Applied Phys.*, Vol. 13, pp. 526–569, September 1942.
38. M. A. TUVE, L. R. HAFSTAD, and O. DAHL, " High-voltage technique for nuclear physics studies," *Phys. Rev.*, Vol. 48, pp. 315–337, August 1935.
39. D. H. SLOAN and E. O. LAWRENCE, " Production of heavy high-speed ions without the use of high voltages," *Phys. Rev.*, Vol. 38, pp. 2021–2032, December 1931.
40. D. H. SLOAN and W. M. COATES, " Recent advances in the production of high-

speed ions without the use of high voltages," *Phys. Rev.*, Vol. 46, pp. 539–542, October 1934.

41. E. O. LAWRENCE and M. S. LIVINGSTON, "The production of high-speed light ions without the use of high voltages," *Phys. Rev.*, Vol. 40, pp. 19–35, April 1932.
42. M. S. LIVINGSTON, "The magnetic resonance accelerator," *Rev. Sci. Instruments*, Vol. 7, pp. 55–68, January 1936.
43. E. O. LAWRENCE and D. COOKSEY, "On the apparatus for the multiple acceleration of light ions to high speeds," *Phys. Rev.*, Vol. 50, pp. 1131–1140, December 1936.
44. E. O. LAWRENCE, L. W. ALVAREZ, W. M. BROBECK, D. COOKSEY, D. R. CORSON, W. W. SALISBURY, and R. L. THORNTON, "Initial performance of the 60-inch cyclotron of the William H. Crocker Radiation Laboratories, University of California," *Phys. Rev.*, Vol. 56, p. 124, July 1939.
45. R. WIDEROE, "New principle for production of high potentials," *Arch. Elektrotech.*, Vol. 21, pp. 387–406, December 1928.
46. D. W. KERST, "The acceleration of electrons by magnetic induction," *Phys. Rev.*, Vol. 60, pp. 47–53, July 1941.
47. D. W. KERST, "20-million electron-volt betatron or induction accelerator," *Rev. Sci. Instruments*, Vol. 13, pp. 387–394, September 1942.

CHAPTER 3

ELECTRON MICROSCOPES

3·1. Reason for Electron Microscope. The objectives pursued in the historical development of the electron microscope have been twofold. On the one hand, there has been the desire to study the processes of electron emission by the detailed examination of emitting surfaces; on the other, it was recognized that electrons offered a means of overcoming the limitations of the light microscope in the direct observation of very small structures of an arbitrary sort. It is this latter circumstance which has given the electron microscope vital importance in all fields of scientific endeavor as well as in many branches of industry and technology. It will, thus, be discussed immediately. The *emission microscopes* form the subject matter of the last section of this chapter.

The unaided human eye, owing to the coarseness of the structure of its retina as well as to shortcomings of its lens action, is normally unable to distinguish, as separate, points or lines lying much less than 0.1 millimeter apart. Comfortable observation requires that these separations should be at least 0.2 millimeter. The purpose of the microscope is to form images of objects which are magnified to such an extent that the eye can recognize in them object structures involving much finer details. Thus, if the details of the interesting structure involve separations of 2 microns ($2 \cdot 10^{-3}$ millimeter), it will be necessary to employ a microscope with a magnification of at least 100 to render the details easily visible. Generally, as the magnification is increased, more and more fine structure appears. However, it was soon found that this process cannot be carried on indefinitely. In particular, for a given light-microscope objective, if the magnification of the eyepiece or if the tube length is carried beyond a certain point, no new structure is revealed; the image simply becomes larger and all outlines become increasingly diffuse in proportion to the enlargement.

During the latter part of last century Ernst Abbe was able to show that this circumstance is not necessarily due to the optical imperfection of the lens system employed. It is, rather, inherent in the imaging medium. If a single light-emitting point is imaged by a perfect lens intercepting a pencil of light rays of half-angle (*aperture*) α (Fig. 3·1), its image is not a point, but a disk of light (*Airy disk*) whose intensity drops

to zero at a distance

$$M \cdot d = \frac{0.61\lambda \cdot M}{\sin \alpha} \qquad [3\cdot1]$$

from the center. Here λ is the wave length of the light with which the image is formed and M the magnification of the image. For two points separated a distance smaller than d the light distributions coalesce in such a manner that the observer is likely to conclude that only a single source is present. If the two points are separated by more than the distance d, on the other hand, they will be recognized in the image as two

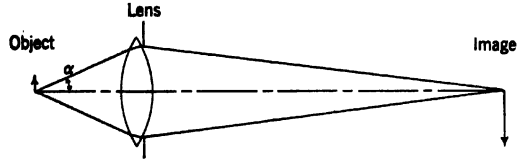


FIG. 3·1. The Angular Aperture of a Lens.

sources. Thus d is taken as a measure of the resolving power of an ideal lens with aperture α. The same conclusion is reached for an object which does not emit light itself, but is illuminated by an external source. The basic phenomenon in either case is the diffraction of the light at the aperture of the lens, causing it to deviate from the laws of geometrical optics; the latter, for a perfect lens, would require a point image for a point source. To a very minor degree the numerical factor in Eq. 3·1 is dependent on the character of the structure observed and the arrangement of the illumination.

If the object which is observed is immersed in a medium of index n relative to the medium in image space, the limit of resolution must be generalized to take the form

$$d = \frac{0.61\lambda}{n \sin \alpha} \qquad [3\cdot2]$$

$n \sin \alpha$ is commonly designated by the term *numerical aperture* (N.A.).

The maximum useful magnification M_u of an image formed by an objective of given numerical aperture with radiation of given wave length is obtained by setting $Md = 0.2$ mm $= 0.02$ cm, the limit of resolution (for comfortable viewing) of the average human eye, and

solving for M:[1]

$$M_u = \frac{0.033n \sin \alpha}{\lambda} \qquad [3\cdot3]$$

The quality of good modern microscope objectives is such that the theoretical limits of the resolving power d and useful magnification M_u are practically attained, even for apertures in excess of 70 degrees, for which sin α approaches unity very closely. The eyepiece, or any other device for increasing the magnification beyond that of the objective alone, need never be a factor limiting the resolving power. The imaging pencils leaving the objective are normally so narrow that neither further diffraction effects nor lens aberrations can easily become significant.

Equations 3·2 and 3·3 indicate that for optimum resolution λ should be made as small as possible and α and n as large as possible. Although available liquid immersion media have values of n approaching 1.7, the more convenient ones, cedar and paraffin oil, come closer to $n = 1.5$. Accordingly, the highest numerical aperture of commercial light-microscope objectives is N.A. = 1.4. For visible light, furthermore, the shortest wave lengths are of the order of 0.4 micron = 4000 Angstrom Units (A.U.), corresponding to the bluish violet. Higher resolution can be attained by passing beyond the visible range, into the ultraviolet. Here, however, a limit is soon reached, since materials suitable for making precision lenses are not adequately transparent below 2000 A.U. Above this value quartz optics are employed.

From the above facts the following limits can be deduced for the resolution and useful magnification of the light microscope:

1. For direct visual observation with $\lambda = 4200$ A.U. and a numerical aperture 1.4:

$$d = 1800 \text{ A.U.}, \qquad M_u = 1100$$

2. For microphotography in the ultraviolet with the mercury resonance line $\lambda = 2537$ A.U. and a numerical aperture 1.4:

$$d = 1100 \text{ A.U.}, \qquad M_u = 1800$$

In round figures, thus, the limit of resolution of the light microscope is 1000 A.U. and the maximum useful magnification, 2000.

[1] Here, as in all other equations involving dimensional coefficients, the several quantities are measured (unless specified otherwise) in the units of the c.g.s. (centimeter-gram-second) system. Lengths, such as λ, are measured in centimeters, masses in grams, and times in seconds. Electrical and magnetic quantities are given in practical units: potentials in volts, currents in amperes, magnetic fields in gauss, etc.

The study of detail finer than 1000 A.U. has necessitated in the past the employment of indirect physical methods, such as centrifuging, sedimentation, filtration, and x-ray and electron diffraction, which, in general, give information of only statistical character. The determination of details of shape and structure of individual units requires direct observation with improved resolution.

The only possible way of increasing the resolving power is seen to lie in the use of radiations of shorter wave length. The application of x-rays, with wave lengths of the order of 1 A.U., is here excluded by the fact that for x-rays refractive media with indices of refraction differing from unity by an adequate amount are not known and, hence, x-ray lenses do not exist. On the other hand, the wave lengths associated with high-speed electrons are known to be short even compared to those of ordinary x-rays. As Louis de Broglie first showed, they are given by

$$\lambda = \frac{12.3}{V^{1/2}} \text{ A.U.} \qquad [3\cdot4]$$

V being the potential difference in volts through which the electrons have been accelerated.[2] Thus for $V = 50{,}000$ volts, $\lambda = 0.0535$ A.U. At the same time, as has already been seen, both electric and magnetic fields may act as electron lenses, so that the basic requirements for the construction of an electron microscope exist. If electron lenses could be made as nearly perfect, that is, as free from aberrations, as light lenses, a resolution far beyond atomic dimensions ($\sim$1 A.U.) could be attained. At the present time this limit is scarcely approached. Aberrations inherent in the type of electron-optical systems employed prescribe upper limits to the size of the effective apertures which may be used. These lie in the range between 0.01 and 0.001 radian. This leads to a practical limit of resolution of the electron microscope, which is of the order of 10 A.U. Even this signifies a gain by a factor of 100 in resolving power and useful magnification as compared with the light microscope.

Table I indicates the methods of investigation suitable for the study of objects having dimensions of different orders of magnitude as well as the resolving powers and useful magnifications required in direct observation. The examples are taken largely from the field of biology. Even now objects in the range from the internal structure of bacteria

[2] Equation 3·4 applies for relatively low accelerating voltages ($\lesssim$20,000 volts). A more generally valid expression for the electron wave length is

$$\lambda = \frac{12.3}{(V + V^2 \cdot 10^{-6})^{1/2}} \text{ A.U.}$$

TABLE I

Object to Be Investigated	Resolving Power Necessary for the Resolution of Objects Separated by Their Diameters (Microns)	Magnification Necessary to Readily Distinguish Objects Visually	Means for Investigating and Obtaining Enlarged Images
Ordinary objects	....	1	Eye
Fine machine work Flaws in jewels	25–100	8	Magnifying glass
Pond life Fungi	10–25	20	Low-power compound light microscope
Bacteria	1–2	200	Medium-power compound light microscope
Structure of bacteria	0.25	800	Electron microscope High-power compound light microscope using oil-immersion objectives and ordinary or ultraviolet light
Large viruses	0.10	2000	Electron microscope High-power compound light microscope using ultraviolet light
Colloidal particles	0.05	4000	Electron microscope
Small viruses Large molecules	0.01	20,000	Electron microscope Ultracentrifuge
Small molecules	0.002	100,000	Electron microscope Analytical chemistry Electron and x-ray diffraction
Atoms	0.0001	2,000,000	Spectroscopy X-ray and electron scattering Other methods of atomic physics

down to molecules of ordinary size are seen to be the special province of the electron microscope. Structures of comparable dimensions play a significant role in many phases of chemistry and metallurgy, giving this instrument equal importance in these fields.

3·2. Magnetic Microscope; Light-Optical Analogue. Among the different types of electron microscopes those employing magnetic fields for the formation of the electron image have proved very successful. The general construction of this type of instrument is shown, compared with that of a light microscope, in Fig. 3·2. In both cases radial separations have been exaggerated in comparison with axial distances, so as to represent more clearly the ray paths. In the magnetic microscope the aperture angles of the imaging pencils have been greatly exaggerated for the same reason.

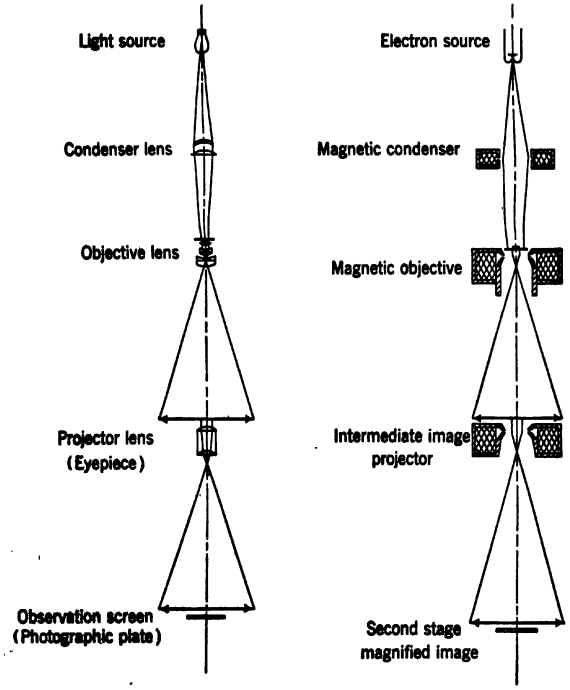


FIG. 3·2. Comparison of the Arrangement and the Ray Paths in the Light Microscope and the Magnetic Electron Microscope.

Consider, first, the light microscope. The source here consists most commonly of an incandescent lamp, whose filament is imaged by a lens system, the condenser, either on the object itself or, so as to obtain more uniform illumination, into the lower focal plane of the objective. The latter form of illumination (*Koehler illumination*) is shown in the dia-

gram, since it corresponds more closely to the normal condition of operation of the electron microscope. By changing the strength (that is, by adding or withdrawing lens components) or the position of the condenser the convergence of the beams illuminating the individual points of the object may be varied.

In the object, usually mounted on a glass slide, the beams experience a modification by absorption, refraction, scattering, and diffraction. Thus the light pencils leaving the various parts of the object will, in general, differ in intensity, color, and even in divergence from the corresponding illuminating pencils. The former, containing all the information required regarding the variation of the transmission properties of the object over its area, are refracted by the microscope objective so as to form a magnified intermediate image. The rays from the central portion of this image are refracted once more by a second compound glass lens, the eyepiece, forming the final image on the ground-glass focusing screen or photographic plate.

The general arrangement of the magnetic electron microscope is seen to correspond qualitatively to that of the compound light microscope in practically all respects. Here also the imaging rays — electrons in this case — leave a source, which in all the more recent instruments is a heated filament. They are accelerated by the difference in potential between this cathode and an apertured anode and, having passed through the latter, maintain a constant velocity throughout the instrument. A magnetic condenser lens, formed by a coil enclosed in an iron shield with a central ring air gap, regulates the convergence of the electrons on the object; variations in the lens strength of the condenser are obtained by changing the coil current. In Fig. 3·2 the case of a very low condenser current, causing an image of the source to be formed at considerable distance below the object, is illustrated. Under these circumstances the convergence angle of the pencil illuminating a given object point is very small.

As in the light microscope, the object modifies the illuminating pencils passing through it. In this modification, however, only scattering and diffraction by the object structure play an appreciable role, the degree of scattering increasing with the object thickness and density. The rays scattered through an angle larger than the effective aperture of the objective, which lies in the neighborhood of $3 \cdot 10^{-3}$ radian, either do not enter the objective at all, or, because of the aberrations of this lens, contribute merely to the brightness of the background of the image.[3] Since all matter is quite opaque to electrons, the specimens, which themselves are only 100 to 1000 A.U. thick, are not mounted on ordinary glass

[3] See section 19·4.

slides, but, instead, on membranes only about 100 A.U. thick, or are supported in space by their own cohesive forces.

The intensity in the central pencils leaving the object points is a measure of the transmission properties of the corresponding portions of the object. The term *central pencil* here refers to the bundle of electron rays, with an angular aperture equal to the effective aperture of the objective, which passes through the objective aperture and contributes to the formation of the image. These central pencils are refracted by the objective so as to form a magnified intermediate image, as in the light microscope. Like the condenser, the objective is a magnetic lens. The principal difference between them lies in the pole-piece structure which is designed to produce a lens of very short focal length.

The central part of the intermediate image is usually selected by an aperture in a disk and is greatly magnified by the last magnetic lens, the projector, which corresponds to the eyepiece of the light microscope. The disk is usually coated with fluorescent material which facilitates adjustment by rendering the intermediate image visible. The construction of the projector is similar to that of the objective. As in the condenser lens, the strength of both these lenses is varied by changing the coil currents. In the objective lens such a variation is employed to focus the image. At the intermediate image the aperture of the imaging pencils is so small that a change in the strength of the projector lens does not affect the sharpness, but merely the size of the image.[4] Hence the current adjustment of the projector is employed simply to control the magnification of the final image.

TABLE II

DIFFERENCES BETWEEN THE LIGHT MICROSCOPE AND A TYPICAL MAGNETIC ELECTRON MICROSCOPE

	Light Microscope	Magnetic Electron Microscope
Image-forming radiation	Light	Electrons
Medium of travel	Air	Vacuum ($\sim 10^{-4}$ mm Hg)
Nature of lenses	Glass	Magnetic fields
Object mounting	Glass slides	Thin films ($\sim 10^{-6}$ cm)
Main source of contrast	Absorption	Scattering
Focusing	Mechanical	Electrical (changing objective coil current)
Adjusting magnification	Exchanging lenses	Changing projector coil current

Figure 3·2 adequately demonstrated the similarities of the light microscope and the magnetic electron microscope. So as to balance the representation, Table II lists the more obvious differences of these two instru-

[4] See section 4·5.

ments. Whereas the analogy in the function of the several parts of the two microscopes is almost perfect, both the construction and the method of operation of the instruments differ radically.

The design and construction of magnetic electron microscopes are discussed in greater detail in the two succeeding chapters. The next four sections are devoted instead to a description of less widely used electron microscopes.

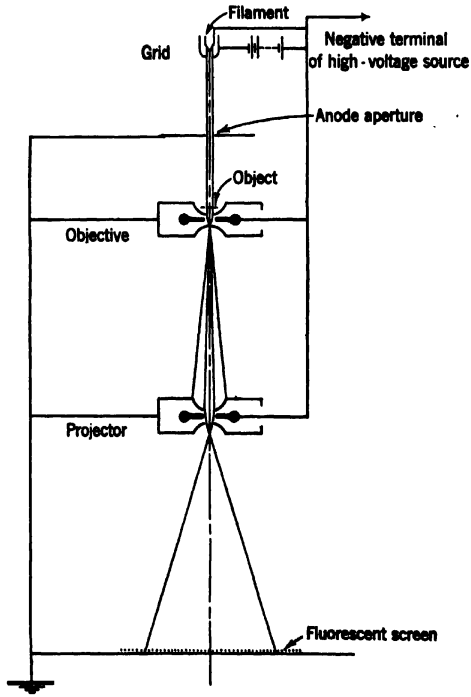


FIG. 3·3. The Electrostatic Microscope. (Schematic Diagram.) (Mahl.)

3·3. The Electrostatic Microscope. The electrostatic electron microscope (Fig. 3·3)[5] resembles the magnetic microscope both in general arrangement and in the process of image formation. The primary difference resides in the nature of the objective and projector, which are here electrostatic lenses.

In the simple model diagrammed in Fig. 3·3 the electron source, a

[5] See Mahl, references 1 and 2.

tungsten filament, is surrounded by an apertured grid, normally operated at a potential about 200 volts negative with respect to the cathode. This electrode serves to control the divergence and the intensity of the electron beam. After having been accelerated to, and through, the anode aperture, at ground potential and 40 to 60 kilovolts positive with respect to the cathode, the electrons strike the object. The modified electron beam then enters the objective, which is here shown as an electrostatic unipotential lens, and is refracted so as to form the intermediate image. The central portion of the intermediate image is again selected by an aperture and is magnified once more, to form the final image on a fluorescent screen or photographic plate, by a projector lens similar to the objective.

The objective and the projector lenses consist, essentially, of a three-aperture system, of which the outer two apertures are connected to ground, the central one to the cathode. Before their construction is discussed in greater detail, it is well to examine the reasons for selecting a unipotential lens with negative center electrode. In principle, immersion lenses, that is, lenses for which the velocities of entrance and exit of the electrons differ, or unipotential lenses with positive electrode, could serve equally well. The last alternative is rendered impractical by the fact that very much greater voltage differences between electrodes are required to obtain a given refractive power or focal length in a lens with positive center electrode than in one with negative center electrode.[6] Immersion lenses, on the other hand, would require, if they are to be free from the same objection, the existence of relatively long electron paths at greatly reduced velocity, either between the two lenses or between the projector and the final image. Since slow electrons are more readily deflected by disturbing magnetic fields, this would require an unreasonably perfect degree of shielding. Finally, the employment of unipotential lenses with negative center electrode makes possible the connection of the center electrode to the cathode, so that only a single voltage above ground (with the exception of the battery sources of the grid bias and the filament current) is required in the whole microscope. Since (at least for relatively low voltages) the lens action of electrostatic lenses depends only on the ratios of the electrode potentials, the sharpness and magnification of the image are thus unaffected, to a first approximation, by changes in the high voltage. This minimizes the requirements of voltage constancy.

At first sight it might appear that the same condition could be attained with variable voltage on the several lens electrodes, provided that each voltage is derived from a voltage divider fed by the same source. In

[6] See section 13-6.

fact, this is true only if the source itself is a constant direct-current source and if the ratio of the currents drawn by the several electrodes remains constant. If the source fluctuates in voltage, the unavoidable reactive components of the circuits connecting the electrodes will cause a fluctuation in the instantaneous potential ratios, and hence in the refractive powers of the lenses.

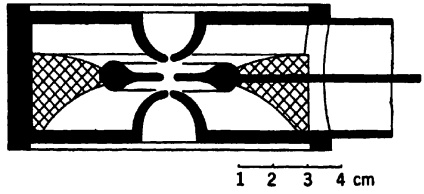


FIG. 3·4. Simple Projector of the Electrostatic Microscope. (Mahl, reference 2. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

To return to the construction of the lenses, Fig. 3·4 shows a schematic diagram of a particular projector lens. Since the limiting factor in the attainment of short focal lengths is electric breakdown between the electrodes, great care must be taken in the selection of the material and in the shaping and polishing of the electrodes. The electrodes shown in Fig. 3·4 were made of chrome-nickel steel, which, according to Gölz,[7] has a breakdown strength of 275 kilovolts per centimeter, even after an earlier breakdown has taken place. For brass the breakdown strength, under similar circumstances, was found to be only 110 kilovolts per centimeter, for aluminum 163 kilovolts per centimeter. The center electrode is toroidal in shape and mounted on a hard-rubber ring, the high-voltage lead being brought out laterally. In the objective the bottom electrode is given an aperture about 0.1 millimeter in diameter, this acting as the physical objective aperture. As in the magnetic electron microscope[8] the effective aperture depends on the spherical aberration of the objective and on the aperture of the illuminating pencils. In view of the necessity of having the focal plane of the objective sufficiently far above the upper electrode to permit bringing the object in coincidence with it, the minimum objective focal length at 60 kilovolts was found to be, in a particular case,[9] 5 millimeters. For the projector, where this requirement does not exist, and the least focal length is determined exclu-

[7] See reference 3.

[8] See section 19·4.

[9] See Mahl, reference 2.

sively by the closest safe approach of the electrodes, this was found to be 3 millimeters at 60 kilovolts, the corresponding least separation being 2.2 millimeters.

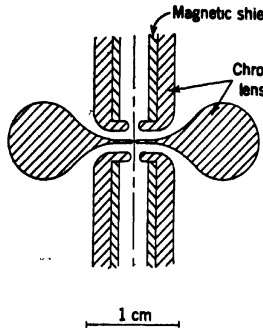


FIG. 3·5. Electrostatic Lens with Narrow Central Aperture. (von Ardenne.)

A different lens design, permitting even shorter focal lengths (2.5 millimeters at 60 kilovolts), but less desirable in a number of other respects, is shown in Fig. 3·5.[10] The negative central electrode, at a potential one-fifth that of the outer electrodes, is provided with an aperture 0.125 millimeter in diameter. Its disadvantages are that it is more readily influenced by voltage fluctuations than the lens with cathode-connected center element, that the beam is peculiarly sensitive to the charging up of contaminations on the central aperture, and

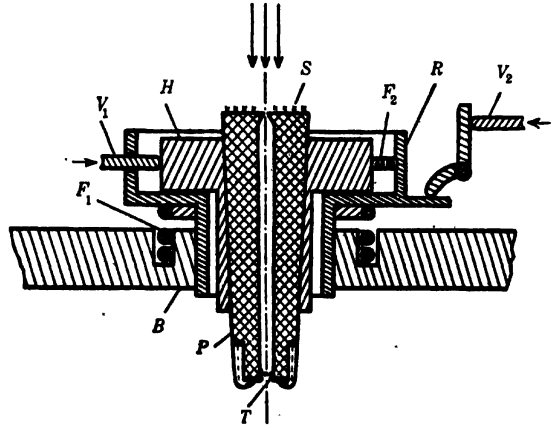


FIG. 3·6. Object Stage of the Electrostatic Microscope. (Mahl, reference 2. By permission of the Alien Property Custodian in the public interest under License No. A-563.)

that the spherical aberration of this lens is much higher than that of a lens with large central aperture.[11]

The entire microscope, as shown in Fig. 3·3, is a fixed-focus, fixed-

[10] See von Ardenne, reference 4.

[11] See section 17·2.

magnification system. The focusing of the object is brought about by mechanical displacement, that is, by moving it into the plane conjugate to the final image plane. Figure 3·6 shows the construction of the object stage, permitting vertical and lateral motion. The conical object cartridge itself has at the top a luminescent screen S for checking the alignment of the electron beam, at the bottom a small apertured disk T, which supports the object. If mounted on a thin film, the latter is placed on the side of the disk away from the objective, as otherwise the film might be pulled off by the strong electrostatic field. The cartridge is introduced into the stage H, which is translated horizontally on the shell R by the screws V_1 acting against spring pressure F_2. The shell R, in turn, slides vertically in the collar B, being pushed down by the action of the screw V_2 and up by the helical spring F_1. Both V_1 and V_2 are actuated by vertical shafts from a point near the viewing port.

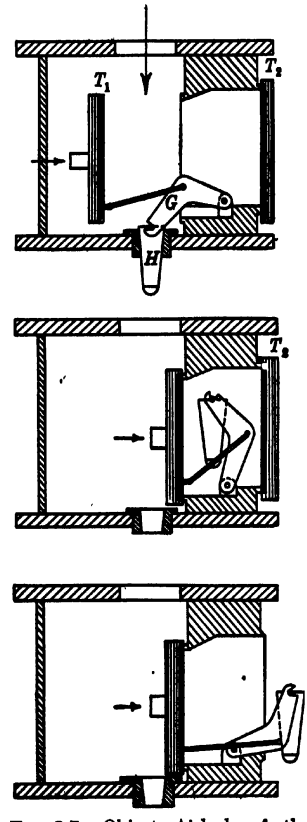


FIG. 3·7. Object Airlock of the Electrostatic Miscroscope. (Brüche and Gölz, *Jahrb. der AEG-Forschung*, Vol. 7, p. 66. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

The introduction of the cartridge into the stage is carried out, in a more recent instrument,[12] through an airlock (Fig. 3·7). After the outer gate T_2 is opened, air entering as the lever is turned, the cartridge H supported by the fork G may be exchanged and then pushed back into the chamber. As the outer gate T_2 is closed, the chamber is automatically connected to the fore vacuum and evacuated. When the inner chamber T_1 is opened, the fork G is pulled over toward the stage and deposits the cartridge therein, making the instrument ready for focusing and observation.

In the same instrument the drawback of a fixed magnification is considerably reduced by providing a double projector lens as shown in Fig. 3·8. The central electrode of either, or of both, of the individual

[12] See Brüche, reference 5.

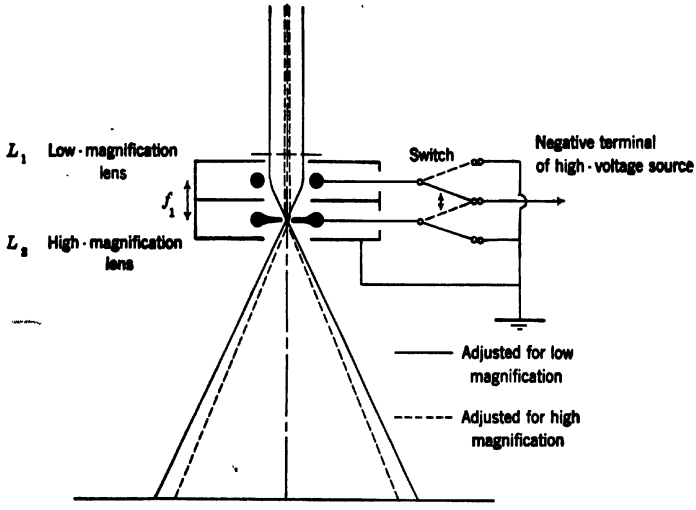


FIG. 3·8. Two-Lens Projector of the Electrostatic Microscope.

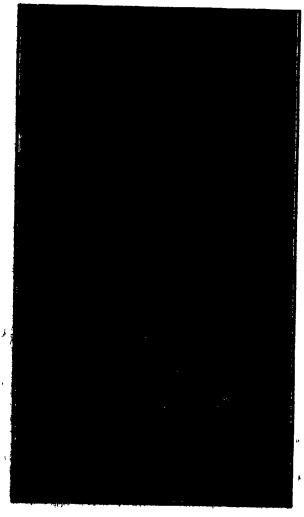


FIG. 3·9. The Electrostatic Microscope of the AEG. (Mahl, *Metallwirtschaft*, Vol. 19, p. 488, 1940.)

lenses may be connected to the cathode, the remaining electrodes being grounded. The upper, long-focus, lens forms a large-diameter image with a magnification of about 1000, serving as a survey image. Since the lower lens, with its relatively narrow aperture, is placed in the focal plane of the upper lens, it does not restrict the field imaged by the latter. The lower lens, acting alone, gives a high-magnification image (magnification 10,000) with a smaller object field. The two lenses acting simultaneously yield an intermediate magnification.

The entire instrument, whose column is carefully shielded magnetically with mu-metal, while the lenses are encased in iron, is shown, with its accessory equipment, in Fig. 3·9.

A still more recent electrostatic electron microscope, making use of the same

type of lenses as Mahl's instrument, is the G.E. Simplified Electron Microscope, shown in Fig. 3·10.[13] To achieve a short microscope column, which is in a horizontal position, three stages of magnification are employed. Thus an electron image with a magnification of 500 to 1000 is formed on the final fluorescent screen. The screen is viewed with a magnifier enlarging the image eight times, yielding a total magnification of 4000 to 8000. The application of this fourth, light-optical, stage of enlargement has the advantage of increasing the apparent size

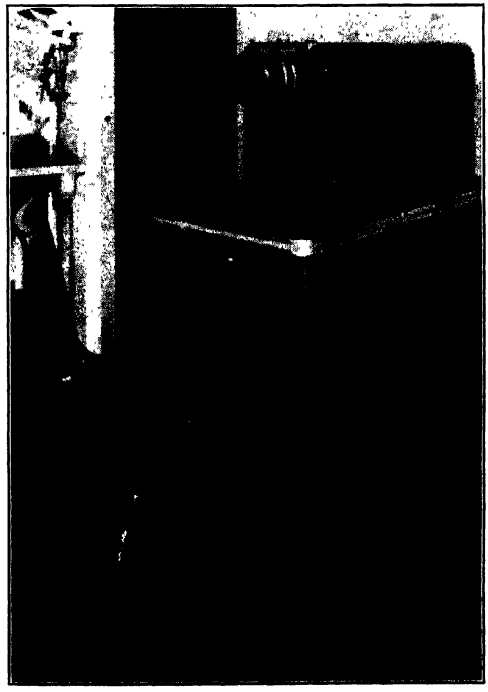


FIG. 3·10(*a*). Simplified Electron Microscope of the General Electric Company, External Appearance. (Courtesy *J. Applied Phys.*, reference 6.)

of the image without correspondingly decreasing its brightness. Focusing is accomplished by displacing the specimen along the axis of the instrument. For permanent record, the image on the fluorescent screen is photographed with a camera. For screens which permit an enlargement of 10 to 20 the maximum useful magnification of the instrument

[13] See Bachman and Ramo, reference 6.

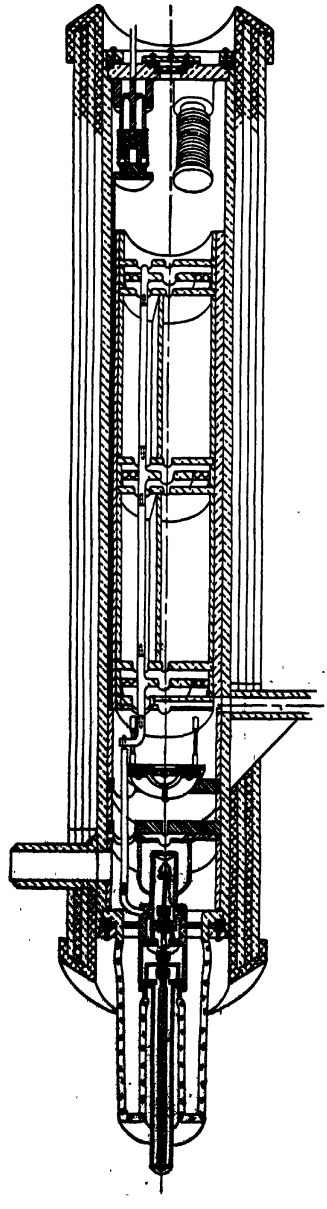


FIG. 3-10(b). Simplified Electron Microscope of the General Electric Company. Section through Electron-Optical System. (Courtesy *J. Applied Phys.*, reference 6.)

thus becomes 10,000 to 20,000; for higher magnifications the limited resolution of the screen impairs the quality of the image. The lower efficiency of the indirect method of recording the images causes the exposure times to be longer than for microscopes in which the electrons fall directly on the photographic plate. Exposures range from half a minute upwards.

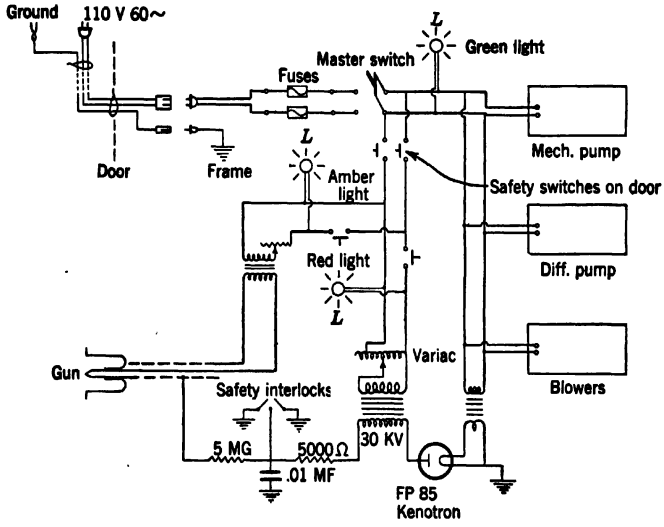


FIG. 3·11. Complete Wiring Diagram of the General Electric Electron Microscope. (Courtesy General Electric Company, Schenectady, N. Y.)

The wiring diagram of the 30-kilovolt supply as well as of the pumping system is shown in Fig. 3·11. All this equipment is contained in the cabinet back of the microscope column. The microscope is evacuated by an air-cooled diffusion pump backed by a rotary oil pump. The high voltage consists of the rectified output of a 30-kilovolt transformer smoothed by a single 0.01-microfarad condenser.

As compared with the magnetic microscope, the primary advantage of the electrostatic instrument is simplicity in the electrical equipment. On the one hand, the current sources for the lenses are omitted entirely; on the other, the requirements on the constancy of the operating voltage are greatly reduced. The use of unrectified alternating voltage, produced directly by a high-voltage transformer, is, it is true, inadmissible unless, by the employment of a strongly negative grid, the fraction of the cycle utilized for imaging is reduced to a very small quantity.

Without this precaution the voltage fluctuations will be reflected in an oscillation of the image, brought about by the varying deflection of the beam by the imperfectly shielded earth's field. Although this effect, too, may be reduced by field-compensating coils and special care in magnetic shielding, it is more convenient to filter the high voltage so that the remaining ripple is less than 1 per cent. This order of voltage constancy is also required by the relativistic aberration of the electrostatic lenses.[14] Compared with the constancy demanded by the magnetic microscope (~0.01 per cent), this requirement is still very moderate.

The simplicity of the electrical equipment is obtained at the expense of the following factors:

1. A somewhat greater complexity of mechanical design brought about by the necessity of mechanical focusing of the object.
2. The necessity of introducing high voltages at several points on the microscope.
3. A restriction to a few fixed values of the magnification and to lower magnifications, for a given number of lenses and for given overall length of the microscope. For the same voltage, electrostatic lenses cannot be made of as short focal length as magnetic lenses.
4. The impossibility of adaptation for work with very high voltages.
5. An ultimate least-resolvable distance, which is probably larger than for magnetic lenses. Both the chromatic aberration[15] and the spherical aberration[16] of the lenses employed are larger than those of the magnetic lenses customarily used.[17]

Thus it would appear that, unless further developments remove some of these drawbacks,[18] the electrostatic microscope is not capable of the same ultimate performance as the standard magnetic instrument.

3·4. Scanning Microscope. Both the magnetic and electrostatic microscopes form images essentially in the same manner as all familiar light-optical instruments. The complete image of the object is projected instantaneously, with the aid of lenses, onto a viewing screen or a photographic plate. The scanning microscope adopts an entirely different technique, resembling that of television (Fig. 2·2). The object,

[14] See section 18·2.

[15] See section 17·6.

[16] See section 17·2.

[17] Mahl (reference 2) reports an optimum resolution of 80 A.U., but does not regard this as the ultimate limit of the instrument.

[18] As, e.g., by the employment of an electron mirror in place of the objective lens in order to achieve correction of spherical and/or chromatic aberration.

which may be regarded as subdivided into small squares whose side is equal in length to the limit of resolution of the instrument, is scanned by a fine electron beam with a cross-section area approximately equal to the area of one of the elements (Fig. 3·12). As the beam sweeps over a particular element, the element emits or transmits a number of electrons, depending upon the physical properties of this element. These electrons may, in a number of ways, be employed to form, in synchronism with the beam scanning the object, a magnified image of the object.

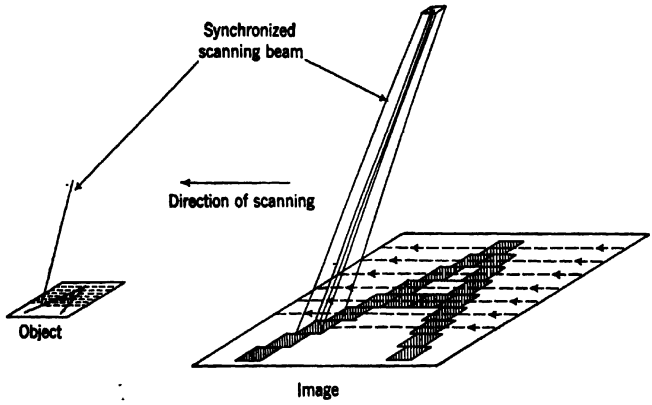


FIG. 3·12. Picture Analysis and Recomposition by Scanning.

A simple form of scanning microscope, borrowed in many respects directly from television practice, is shown in Fig. 3·13. The object-scanning beam, or electron probe, is the reduced image of the crossover of an electron gun formed by an inverted magnetic electron microscope, the reduction ratio being thus given directly by the reciprocal of the magnification of the microscope. A magnetic deflecting field just above the objective of this microscope — the final reducing lens — causes the beam to scan the object in two mutually perpendicular directions. The secondary electrons emitted by the object under the impact of the primary electrons of the probe constitute the signal current. They are drawn to a collecting electrode and generate a signal voltage across a resistor connecting this electrode to ground. This voltage, after amplification, modulates the beam current of a viewing tube, whose screen is scanned in synchronism with the object. If the rate of scanning is fast enough, for example, 30 pictures per second, an observer will, owing to the

persistence of vision,[19] see a complete image of the object on the tube screen, even though only a single image element is excited by the beam at any one time. The magnification of the image is then equal to the ratio of the scanning amplitudes on the screen and on the object. The resolution, on the other hand, is limited by the diameter of the scanning probe and the ratio of the length of a side of the scanning pattern to the number of scanning lines. Normally, these two quantities are made approximately equal.

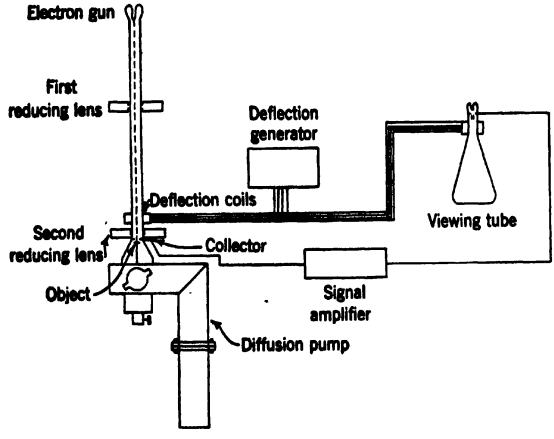


FIG. 3·13. Scanning Microscope for the Direct Viewing of the Image. (Courtesy *ASTM Bull.*, reference 9.)

It is found that with the system described the optimum resolution that can be obtained is of the order of one micron. If the probe is made materially less than a micron in diameter and the amplifier gain is increased to maintain the signal output at the same level, the image on the screen no longer shows any picture detail, but merely random fluctuations in brightness. The *noise*, that is, the random fluctuations in the voltage modulating the beam current, exceeds the amplified signal voltage. This is due in part to the extremely small amount of charge emitted per picture element in consequence of the discreteness of the elements of charge. In addition to this, noise is introduced by the unfavorable conditions of amplification.[20]

Von Ardenne[21] was able to eliminate both sources of fluctuation in a

[19] This may be aided by the persistence of the light emission of the fluorescent screen.

[20] See Appendix I.

[21] See reference 7.

simple manner in a scanning microscope designed specifically for the study of objects by transmission. The noise introduced by the coupling resistor and the amplifier is nullified by dispensing with amplification. At the same time, the *shot noise*, due to the discreteness of the electronic charge, is rendered harmless by increasing the recording period to 10 minutes or more, so that the number of electrons per picture element is large even if the individual element has dimensions comparable with the optimum resolution of the microscope.

In the von Ardenne instrument the image is registered photographically on a recording drum located below the object, the scanning beam itself forming the record, as shown schematically in Fig. 3·14. The probe is formed as in the instrument previously described (Fig. 3·13) by an inverted magnetic electron microscope and is deflected by a pair of coils above the final lens. The currents for these two deflecting coils, determining the displacement of the probe along and normal to the scanning lines, are derived from slide-wire bridges mechanically coupled to the recording drum, several seconds being required for one revolution of the drum. The photographic paper attached to the surface of the drum is displaced about 2500 times as rapidly as the probe moves across the object. The ratio of the two speeds of displacement (for example, 2500) is equal to the magnification of the image. Hence the spreading of the beam after leaving the object is much less than one image element and only serves to fill up to some extent the space between the scanning lines without detracting from the resolution of the image. In order to obtain normal image contrasts, it is necessary to insert an aperture, holding back the more strongly scattered electrons, between the object and the recording drum.

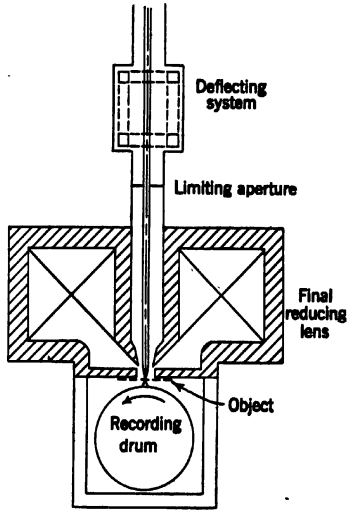


FIG. 3·14. Recording System of von Ardenne's Transmission Scanning Microscope (Schematic).

The ultimate limit of resolution of this scanning microscope for transmission is identical with that of the standard magnetic microscope. It has the advantage over the latter that its resolution is not affected by the velocity losses of the electron within the object, which, in the usual

instrument, give rise to chromatic aberration.[22] This advantage, appreciable only with relatively thick specimens, is scarcely sufficient to outweigh the impossibility of direct observation, the greater difficulties of adjustment, and the longer time required for obtaining an image.

The situation is different, however, if the specimens to be studied are opaque to electrons. The direct observation of such specimens with the standard electron microscope, modified so as to utilize electrons reflected by the specimen surface, is possible only at the expense of a great loss in resolution.[23] Since direct registration of the image cannot be accomplished with opaque specimens in the scanning microscope, a new method must be employed to reduce the objectionable voltage fluctuations superposed on the signal. On the one hand, it is necessary, again, to increase the time of registration. Since this makes impossible the direct observation of the image on the screen, as shown in Fig. 3·13, the viewing tube is conveniently replaced by a facsimile receiver, such as is employed for recording pictures transmitted by radio or telegraph.[24] On the other hand, the thermal noise introduced by the resistance coupling of the secondary-emission collector to the amplifier may be made innocuous by preamplifying the secondary-emission current from the object with the aid of a secondary-emission multiplier.

Figure 3·15 shows the general plan of a scanning microscope for surface observation, which incorporates both improvements.[25] The electron source, at the bottom of the microscope, is a thin tungsten filament. The electron beam leaving it passes through a grid aperture and is modulated by a 3000-cycle-per-second square wave applied to the grid. Beyond this, in passing through the anode, it is accelerated through a drop of 10,000 volts, which represents the beam potential throughout the major part of the instrument. Two electrostatic lenses with negative center electrodes serve to form a greatly reduced image of the source on the object, which is maintained normally at about 800 volts, this relatively low voltage being favorable for obtaining a large difference in the secondary-emission currents from different parts of the object. The limiting aperture of the final lens is located between the two reducing lenses and just below an inclined fluorescent screen.

The secondary electrons given off by the object under the impact of the electron probe are drawn back through the final reducing lens and, being slower than the primary electrons, are focused close to this lens, diverging from this point on. Hence only a small fraction is capable of

[22] See section 19·10*a*.
[23] See section 8·7.
[24] See Young, reference 8.
[25] See Zworykin, Hillier, and Snyder, reference 9.

returning through the limiting aperture; the remainder impinge on the inclined fluorescent screen, giving rise to an emission of light proportional to the strength of the secondary-emission current. This light is collected, to a large extent, by the photocathode of a multiplier (RCA

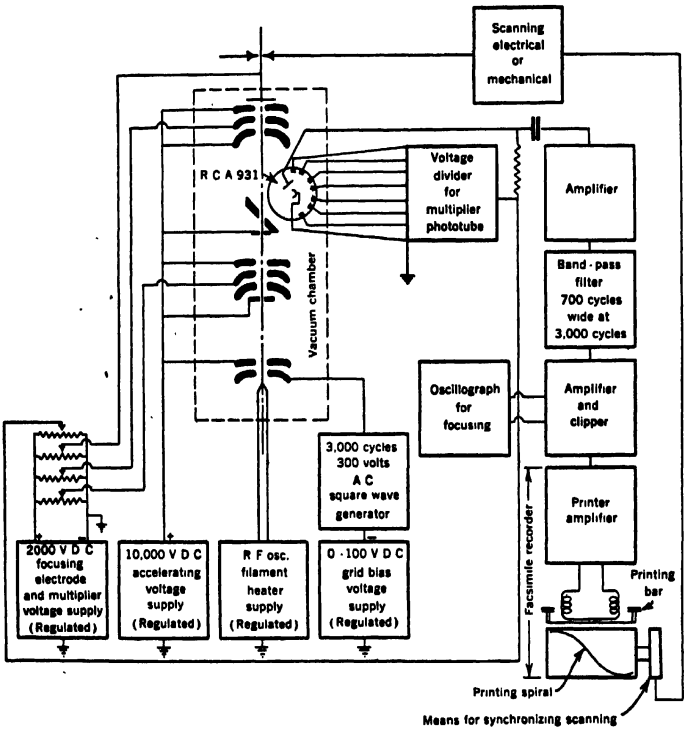


FIG. 3·15. Scanning Microscope for Surface Observation. General Plan. (Courtesy *ASTM Bull.*, reference 9).

931). The output of the multiplier is amplified and filtered and, eventually, *clipped* so as to accentuate the differences in signal between different points of the object, increasing correspondingly the contrasts in the final image. The clipping process is illustrated in Fig. 3·16. The signal is applied to the grid of an amplifier tube which is given a sufficiently strong negative bias that current reaches the plate only during the positive peaks of the signal. The amplified signal, finally, is applied to the

facsimile printer, controlling the pressure of the printer bar against the printing spiral on the rotating drum. In this fashion it determines the density of the impression left on the recording paper along a scanning line. An oscillograph is provided which makes possible the adjustment of the lens voltages for optimum focus, this being indicated by the strongest representation of high-frequency components in the signal. A large scanning spot minimizes the secondary-emission current variations as it passes from one picture element to the next and thus blots out these high-frequency components.

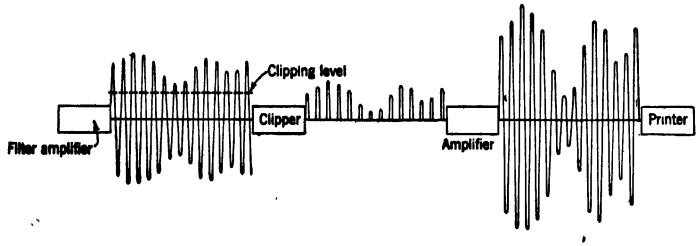


FIG. 3·16. Effect of the " Clipper " on the Signal Applied to the Facsimile Printer.

In the earlier instruments this scanning was carried out mechanically. The displacement of the object was obtained through the use of electromechanical or hydraulic devices and was controlled by the recorder. In a later instrument, shown in schematic cross section in Fig. 3·17, this was replaced by magnetic scanning controlled by the charging of condensers through resistances, the charge and discharge processes being tripped by the recorder. This was found to permit greater precision in the translation of the probe relative to the object than could be attained by any cam-controlled motion of the object itself, with the beam stationary.

To provide space for the deflection coils, as well as for the fluorescent screen and multiplier, which now had to be placed on the object side of the deflection, the probe formed by the original lens is imaged once more, by an added long-focus electrostatic lens, with approximately unity magnification, onto the fixed, relocated object. The deflection coils, wound on a styrol form, are placed just below this lens. Above it is the fluorescent screen with an aperture for the primary beam. Just in front of the object there is an electrode at a low potential which serves to render the returning beam of secondary electrons more diffuse, causing a smaller proportion to strike the aperture in the screen. A large-aperture lens, finally, concentrates the light emitted by the screen on the photo-

cathode of the multiplier, which is located outside the body of the microscope. Since the velocity of the electrons striking the screen corresponds to approximately 9000 volts, it is not difficult, by matching the spectral emission of the fluorescent substance to the spectral sensitivity of the

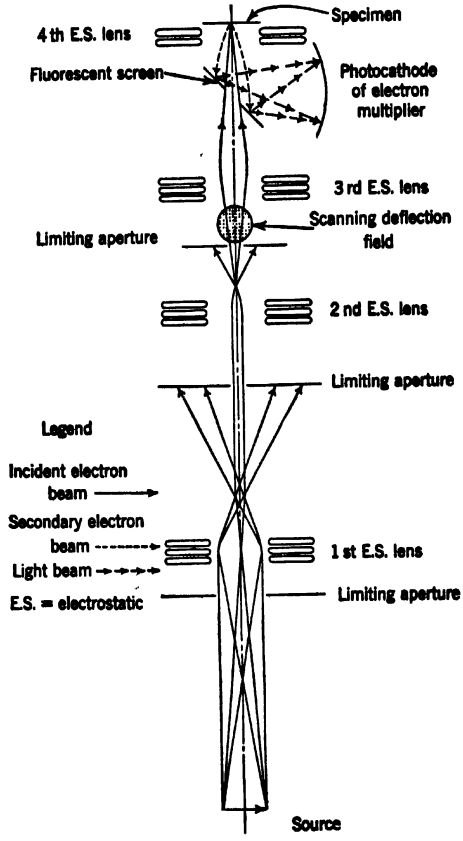


FIG. 3·17. Scanning Microscope for Surface Observation with Magnetic Deflection. Schematic Cross-Section Diagram. (Courtesy *ASTM Bull.*, reference 9.)

multiplier photocathode, to attain a gain in signal as compared with a system in which the secondary electrons from the object fall directly on the first target electrode of the multiplier. Of even greater, and more obvious, value is the simplification in construction which results from the employment of a sealed-off multiplier.

Figure 3·18 shows the general arrangement of the instrument last described. The high-voltage supply is located at the left. In the center is the control panel and the microscope column; the oil diffusion pump is visible below the former. On the table at the right, finally, is the facsimile recorder.

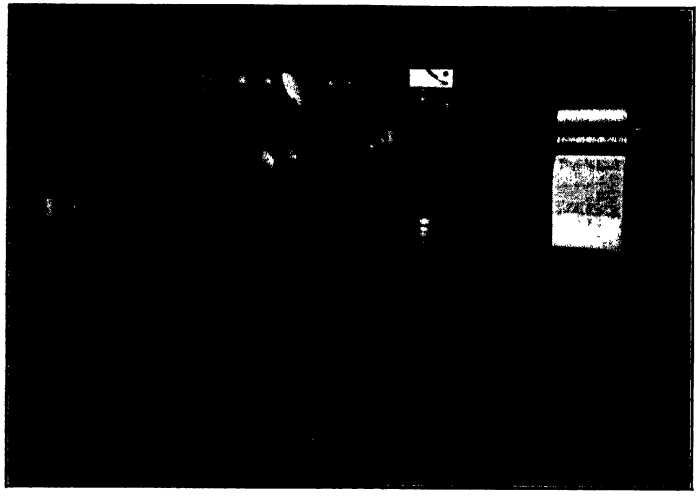


FIG. 3·18. Scanning Microscope for Surface Studies. (RCA Laboratories.)

A scanning micrograph of etched brass, obtained with this device, is shown in Fig. 3·19. The original magnification of this micrograph is 6000. Its resolution may be estimated to be 500 A.U., which must not, however, be regarded as the limit of performance of the instrument. The latter may be expected to lie in the neighborhood of 100 A.U. The fine-line structure of the image can barely be distinguished in the original.

3·5. Point and Shadow Microscopes. The simplest of all possible electron microscopes is realized when a point electron source is placed behind the object, the image being observed on a screen some distance from it. If the path of the rays is rectilinear, that is, if the space considered is substantially field-free or if the direction of the forces acting on the electrons coincides everywhere with their direction of motion, the magnification of the image will be given simply by the ratio of the distance of the screen from the source to that of the object to the source (Fig. 3·20).

Electron sources of the requisite character may be obtained by etch-

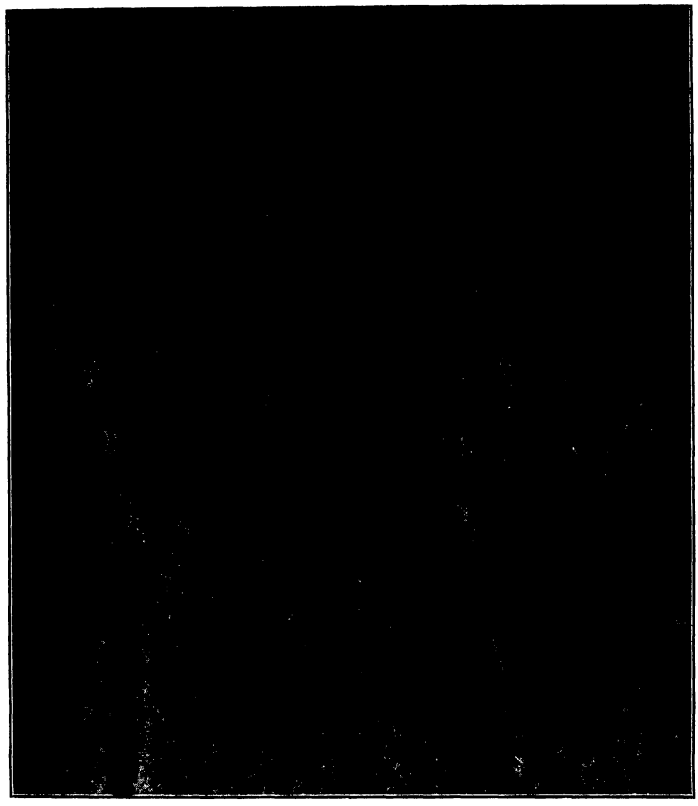


FIG. 3·19. Scanning Micrograph of Etched Brass. Magnification 6000. (Courtesy *ASTM Bull.*, reference 9.)

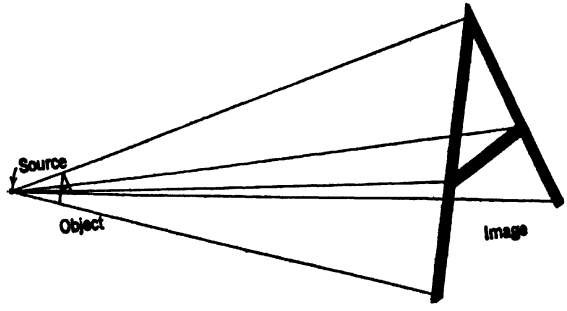


FIG. 3·20. Principle of the Point Microscopes.

ing the end of a tungsten or molybdenum wire in hot fused sodium nitrite and glowing it subsequently in vacuum.[26] In this manner well-rounded points with radii of curvature which are a fraction of a micron can be formed.[27] If a potential difference is applied between the point (as

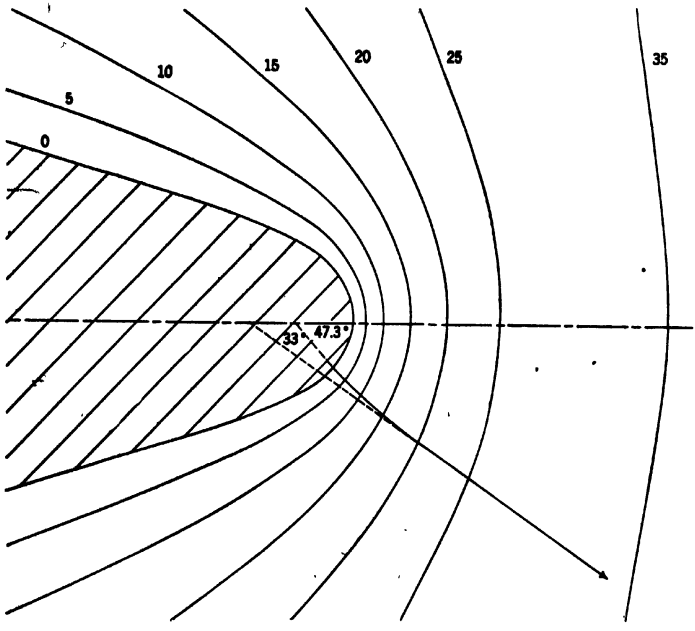


FIG. 3·21. Path of an Electron in the Field Surrounding a Paraboloidal Point.

cathode) and an electrode or screen surrounding it, the field at the surface of the point, being of the order of 10^7 to 10^8 volts per centimeter,[28] is sufficient to draw conduction electrons in large quantities out of the metal. Because of the enormous field at the surface the electrons proceed normal to the surface, attaining, in the case of a spherical point, 90 per cent of their ultimate kinetic energy within a distance equal to nine radii of curvature of the point. With the more or less paraboloidal points normally employed, the paths undergo a small initial curvature, causing them to appear to originate at a point a small distance behind the center of curvature (Fig. 3·21). In either case the electron paths become practically rectilinear within a distance of a few microns.

[26] See Eyring, Makeown, and Millikan, reference 10, and Müller, reference 11.

[27] See Haefer, reference 12.

[28] For a spherical point it is given simply by the ratio of the applied voltage to the radius of curvature.

If, now, an object, largely transparent to electrons of the velocity corresponding to the applied voltage, is placed in front of the point and a fluorescent screen is provided to receive the transmitted electrons, an enlarged image of the object is obtained on the screen. This is the principle of the *point projector microscope* shown in Fig. 3·22.[29] For a thin specimen the contrast is produced by the same processes as in the standard electron microscope. Part of the electrons are scattered, being

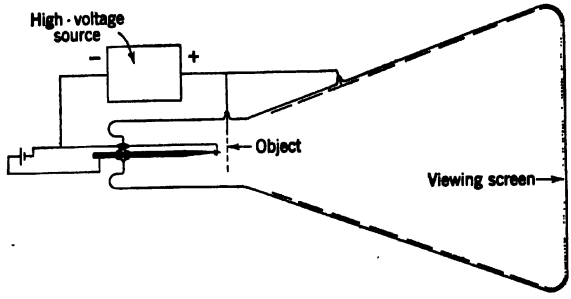


FIG. 3·22. Point Projector Microscope (Schematic).

spread for the most part over the remainder of the image as a diffuse electron fog. For a thick object with clear spaces, a simple shadow projection is obtained, the thicker portions absorbing or reflecting the incident electrons completely, whereas the clear portions let them pass without deflection.[30] The two factors which limit the resolution of the instrument are Fresnel diffraction at the edges of the object on the one hand and the divergence of the imaging pencils due to the small lateral initial velocities of the electrons on the other. As shown in section 19·10*d*, the limit so imposed is of the order of 100 A.U.

In Fig. 3·22 the details of the *object stage* have been omitted. In the actual construction the object mounting was hinged to prevent accidental contact with the point during the processing of the tube. Bimetallic strips and heaters were provided for varying the separation of the object from the point, which was made of especially heavy wire (0.025-inch tungsten or molybdenum) to reduce vibration effects to a minimum. Images of 400-mesh copper screen, obtained with this instrument, varying the object distance, are shown in Fig. 3·23. These were photographed on the screen with an exposure of 1/20 second and an F:4.5 lens. The enormous specific emission of the points, of the order of 10,000 amperes per square centimeter, makes it relatively simple to attain images of considerable brightness even at high magnifications.

[29] See Morton and Ramberg, reference 13.

[30] Except for Fresnel diffraction at the edges.

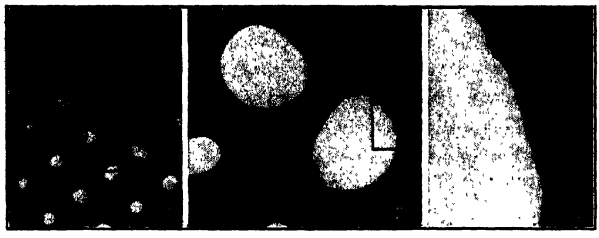


FIG. 3·23. Images of 400-Mesh-per-Inch Copper Screen Formed with Point Projector Microscope, Varying Object Distance. (Courtesy *Phys. Rev.*, reference 13.)

The primary disadvantage of the point projector microscope, attractive because of its extreme simplicity, rests in the great sensitivity of the point cathodes to residual gases in the evacuated tube, making their use impractical in demountable vacuum systems, such as are required for the exchange of the object. The shadow microscope of Boersch[31] (Fig. 3·24) avoids this difficulty, at the expense, it is true, of greatly increasing the complexity of the instrument. Here the point source is formed, in the same manner as the probe in the scanning microscope, by an inverted electron microscope. Electrostatic lenses, similar to those of the previously described electrostatic microscope of Mahl, are employed with a completely unfiltered voltage supply. A negatively biased grid restricts the operation of the microscope to the portions of the cycle during which the applied voltage differs from its maximum value by less than 5 or 10 per cent.

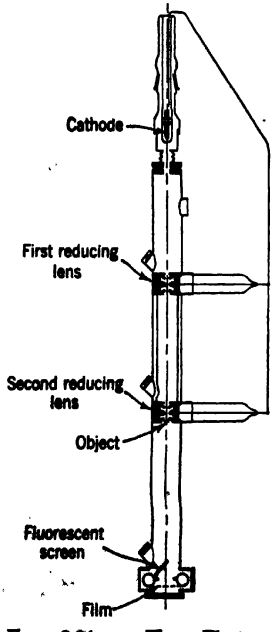


FIG. 3·24. The Electron Shadow Microscope. (Boersch, reference 15.) By permission of the Alien Property Custodian in the public interest under License No. A-563.

The object in its mounting fits into one of the outer electrodes of the final lens, provision being made for lateral displacement in two directions. Furthermore, to vary the magnification, the object can slide up and down in the electrode shell. Focusing is rendered unnecessary by the nature of the instrument. The spherical aberration of the final lens

[31] See Boersch, references 14 and 15.

does not contribute to the unsharpness, but only to the distortion of the image. The presence of diffraction at the object, as well as the requirement that the image detail be not distorted beyond recognition, leads, however, to the same expression for the ultimate resolution of the shadow microscope as for the standard instrument.[32]

In Boersch's instrument the pictures are taken on standard 32-millimeter cine-film. The use of relatively low magnifications (about 1500 or less) is here particularly advisable to attain reasonable exposure times. These lie in general between 1 and 5 seconds. In view of the great sharpness of the pictures (provided that the source-forming inverted microscope is well aligned), they may be subsequently enlarged by a factor of 10 to 20. Boersch has obtained resolutions of 250 A.U. with the shadow microscope.

3·6. X-ray Shadow Microscope. A relatively simple modification of the shadow microscope converts it into a projection microscope utilizing x-rays in place of electrons. It is only necessary to let the sharply focused electron probe impinge on a target, preferably of a heavy metal, so as to render the efficiency of x-ray production a maximum and the penetration of the electrons into the target a minimum. The x-rays formed at the point of impact are then employed for projecting an image of a closely spaced object on a distant screen (Fig. 3·25). Such an instrument has two advantages over the shadow electron microscope. On the one hand it is able to image relatively thick and dense specimens; on the other, the contrasts are almost entirely due to the absorption of the x-rays, so that background fog due to scattering at other points of the image is minimized.

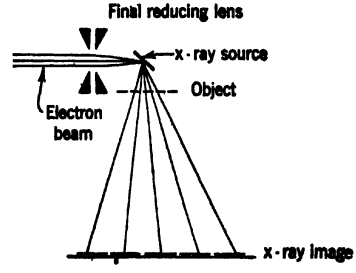


FIG. 3·25. Principle of X-ray Shadow Microscope.

Its disadvantages are the necessity of focusing the probe, as for the scanning microscope, somewhat lower resolution for the same separation between source and object due to the longer wave lengths employed, and, most serious of all, extremely low image intensity. This is due to the low efficiency of conversion of electronic kinetic energy into x-ray energy as well as to the losses arising from the emission of the x-rays in directions which are not utilized. The fact that photographic materials are appreciably less sensitive to x-rays (because of their incomplete

[32] See section 19·10b.

absorption in the emulsion) than to electrons adds to this condition; specially thick emulsions or double-coated films must be rejected because of the resulting loss in resolution. These low x-ray intensities practically exclude the use of the x-ray shadow microscope for obtaining high-resolution, high-magnification images.[33]

3·7. Emission Microscopes. The emission microscopes have the common characteristic that the object constitutes at the same time the source of the electrons, the latter leaving it with energies ranging from zero upward. Historically, they are the earliest of the electron microscopes. It is perhaps most convenient to divide them, on the basis of the origin of the image-forming electrons, into thermionic, photoelectric, secondary-emission, and field-emission microscopes.

[33] An estimate of the relative concentration of energy in an x-ray and an electron shadow image can be obtained quite simply. If the target is solid or thick enough to absorb the electrons fully — a condition which can be shown to make for greatest efficiency — the diameter of the source is equal at least to the depth of penetration of the electrons. Thus the operating voltage V is best selected so that the penetration is equal to the desired resolution d. For tungsten this becomes, according to Eq. 19·38:

$$d = 2{\cdot}10^{-13}V^2$$

Since the efficiency of conversion of electronic energy into x-rays is given, over a wide range, by $1.1{\cdot}10^{-9}ZV = 8{\cdot}10^{-8}V$ (Compton and Allison, reference 16, p. 89), the energy in the x-ray beam is less than that of the electron beam by the factor $0.2d^{\frac{1}{2}}$. Furthermore, this energy is spread over a solid angle 4π, whereas the electron beam is spread over a solid angle $\pi\alpha^2 = \pi d^{\frac{3}{2}}/(Cf)^{\frac{3}{2}}$. As before, the spherical-aberration coefficient Cf will be set equal to 10 cm. Thus for the same magnification and resolution the energy incident on unit area of the x-ray image is less than that incident on unit area of the electron image by the factor $0.01d^{\frac{1}{2}}$. This assumes also equal voltages for the forming of the spots in the two cases. Suppose, now, that the limiting resolution is to be 1 micron and the magnification 100. If it is assumed from Boersch's data (reference 15) that an exposure of 5 seconds suffices to form an image in the electron shadow microscope with a resolution of 250 A.U. and a magnification of 1500, it would follow from Eq. A10 that an exposure of $5/(15{\cdot}40)^2 = 1.4{\cdot}10^{-5}$ second should suffice to form the image with the smaller magnification and resolution. The x-ray energy per unit area of the image being less by a factor of about 10^{-7} (by the formula just derived) than the electron energy per unit area, the exposure required to obtain an x-ray image would be $140{\cdot}S_x/S_1$ second, if S_x and S_1 denote the relative sensitivities of the photographic material to x-ray and electron energy, respectively. The required operating voltage would be about 20,000 volts.

No quantitative information is available regarding the magnitude of the ratio S_x/S_1. It is generally accepted, however, as being considerably less than unity, as may be expected from the more complete absorption of electronic energy in the emulsion. The minimum exposure of the x-ray shadow microscope, even with the moderate resolution of 1 micron, must thus run into several, and probably a very large number of, minutes.

The first application of the thermionic microscope was to the study of oxide cathodes. Both the electrostatic microscope with a single immersion lens, shown in Fig. 3·26, giving magnifications up to nearly 200 with a cathode-to-screen distance of 24 centimeters[34] and the magnetic microscope with either one or two lenses[35] have been extensively used for this purpose. Figure 3·27 shows an early two-lens magnetic microscope. Since the electrons must be accelerated, the magnetic lens action is here supplemented by the electrostatic lens action of the accelerating electrodes. Their shape and arrangement relative to the cathode are of critical importance for the attainment of good images. Both microscopes use accelerating potentials of the order of 1000 volts. Johannson[34] found that in the electrostatic instrument the field aberrations, which are here especially serious,[36] could be reduced materially (increasing the diameter of the useful field 70 per cent) by giving the focusing electrode B_1 a funnel-shaped aperture, as shown in Fig. 3·26. Brüche and Johannson[37] have employed lenses of this type to obtain cinematographic records of the effects of long-time operation and overheating of oxide cathodes. An electron microscope permitting the observation of the cathode in vacuum with either electric, magnetic, or light lenses was constructed by Knecht.[38] The electric and light lenses were attached to a shaft, which could be rotated from the outside, so that one or the other was in place in front of the cathode, the magnetic coils remaining permanently in position. A right-angle prism attached to the light objective

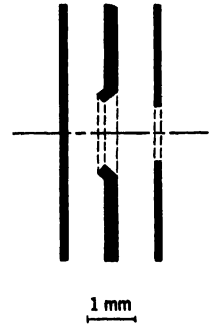


FIG. 3·26. Immersion Objective for Electrostatic Emission Microscope. (Johannson.)



FIG. 3·27. Magnetic Emission Microscope. (Knoll, Houtermans, and Schulze, reference 19.)

[34] See Johannson, reference 17. More recently immersion objectives of the type employed by Johannson, used with accelerating voltages of 20 to 30 kilovolts and with a second stage of magnification, have yielded thermionic-emission pictures with resolutions of the order of 200 to 400 A.U. See Mecklenburg, reference 18.

[35] See Knoll, Houtermans, and Schulze, reference 19.

[36] See sections 16·8 and 17·4.

[37] See reference 20.

[38] See reference 21.

served to make possible observation through a lateral window in the tube.

Knecht found no differences between the magnetic and the electrical images, except for greater unsharpness of the electrical images in the marginal regions. He found, however, like other observers, that much detail, for example, the distribution of adsorbed barium on the surface and the crystal structure of the base metal, was well differentiated in the electron pictures, though invisible or barely indicated in the light pictures.

A wider field was opened to the thermionic microscope when it was realized that the crystalline structure of heated metals could be rendered visible, even at moderate temperatures, by evaporating a thin layer of barium on the surface.[39] This technique has been utilized for following recrystallization processes in iron in passing from the α-phase to the γ-phase, as the temperature is reduced from 700 to 665° C,[40] as well as for a number of other metallurgical processes.[41] The last set of observations was carried out with a single-lens magnetic electron microscope, the magnification being of the order of 20 and the accelerating voltages lying between 3000 and 6000 volts. A similar instrument was constructed by McMillen and Scott for examining ashed biological specimens on a flat platinum electrode.[42]

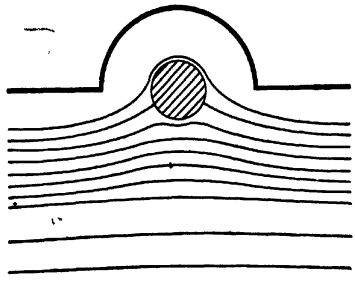


FIG. 3·28. Auxiliary Electrode for Electron-Microscope Observation of Cylindrical Cathodes. (Heinze and Wagener, reference 27.)

All the studies described so far were carried out with flat, generally indirectly heated, cathodes. It is possible, however, to obtain satisfactory images even of wires[43] or cylindrical cathodes[44] with a normal emission microscope. For this purpose an auxiliary electrode at negative potential is placed behind the cathode so as to flatten the equipotential surfaces in front of it. Figure 3.28 shows the arrangement and the potential distribution in the configuration employed by Heinze and Wagener[44] for this purpose.

In addition to these microscopes of more or less conventional type

[39] See Brüche and Johannson, reference 22.
[40] See Brüche and Knecht, reference 23.
[41] See Burgers and Ploos van Amstel, reference 24.
[42] See reference 25.
[43] See Mahl, reference 26.
[44] See reference 27.

must be mentioned the thermionic projection microscopes of R. P. Johnson[45] and S. T. Martin.[46] Johnson's microscope (Fig. 3·29) is especially designed for studying the emission properties of wire surfaces; the wire is stretched on the axis of a cylinder coated with fluorescent material. If the wire is heated and a high potential difference is applied between it and the fluorescent screen, in which a helical anode is embedded, the strong field at the wire surface causes the thermionic electrons to proceed approximately rectilinearly to the screen, forming there an image of the wire surface on which die marks and crystal boundaries are clearly visible. Unlike ordinary microscope images, this image is magnified in one direction only, that is, normal to the wire.

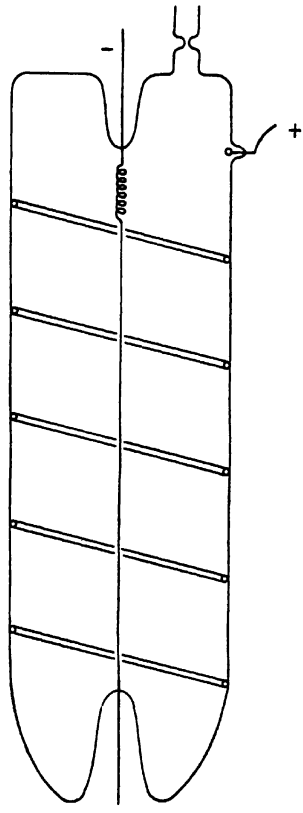


FIG. 3·29. Thermionic Microscope for Observing the Emission from Wires. (R. P. Johnson.)

S. T. Martin has modified Johnson's arrangement by grinding and polishing a single-crystal tungsten sphere and placing it in a spherical envelope coated with luminescent material (Fig. 3·30). The tungsten sphere is heated by electron bombardment from a filament inserted in it. A voltage applied between it and the spherical screen causes thermionic electrons to travel along radial lines to the screen, on which they then produce a pattern indicating the distribution of emission over the surface of the tungsten sphere. Since the work function of the surface of a crystal varies with its orientation with respect to the basic crystal-lattice, these patterns reflect the symmetry properties of the crystal itself. Interesting modifications of the patterns result from the preferential adsorption of contaminants within the tube on differently oriented portions of the surface of the single crystal.

[45] See reference 28.
[46] See reference 29.

The photoelectric microscope serves similar purposes as the thermionic microscope, making use of the fact that all metallic surfaces are photosensitive, to some extent, to ultraviolet light. Frequently the fact that the photoemission is not dependent on the temperature of the specimen represents a material advantage. Although it was realized at a very early stage that photoelectrons could be employed for forming images,[47] Pohl[48] and Mahl and Pohl,[49] employing the instrument shown in Fig. 3·31, may well have obtained the best results in this field. The object, acting as cathode, was illuminated by a quartz-mercury arc through a lateral window. Two diaphragms, as well as the blackening of the inside of the anode cylinder, which surrounds the cathode, prevent an appreciable amount of light from reaching the fluorescent screen at the end of the anode cylinder. To gain sufficient intensity in the image,

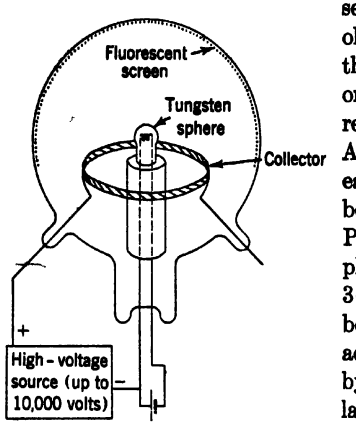


FIG. 3·30. Thermionic Microscope with Radial Projection. (S. T. Martin.)

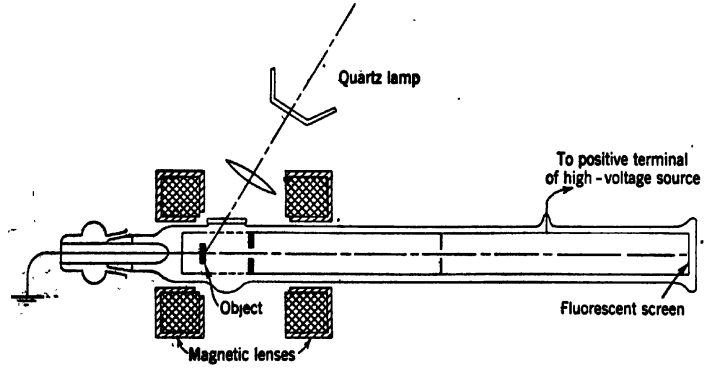


FIG. 3·31. Photoelectric Emission Microscope. (Pohl, reference 32.)

an accelerating voltage of 30 kilovolts was used, the photocurrents having values of 10^{-7} to 10^{-6} ampere. Magnifications ranged up to

[47] See Knoll and Ruska, reference 30, and Brüche, reference 31.
[48] See reference 32.
[49] See reference 33.

about 30. This photoelectric microscope was employed not only for studying changes in crystallization of heated metals, but also for examining effects of surface contaminations, low-temperature melting processes, and semiconducting minerals embedded in Wood's metal.

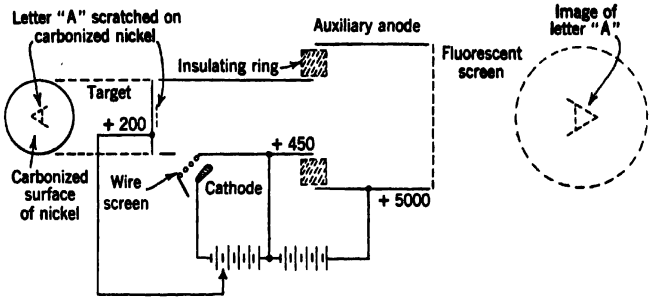


FIG. 3·32. Imaging with Secondary Electrons. (Courtesy *J. Franklin Inst.*, reference 34.)

The possibility of utilizing secondary electrons for forming electron images was first demonstrated by Zworykin[50] with the simple apparatus shown in Fig. 3·32. A simple electrostatic two-cylinder lens images the letter "A" scratched in a metal target, which is bombarded with 20-volt electrons from a thermionic cathode, on a fluorescent screen. A modification of this device, employing a ring filament as electron source so as to make axial symmetry possible and replacing the electrostatic lens by magnetic lenses, has been described by Meschter.[51]

Very interesting pictures of metal foils, in part disintegrating under bombardment, have been obtained by Behne[52] with secondary electrons. Behne employed the secondary electrons given off on the side opposite to that on which the primary electrons are incident. For certain critical velocities of the primary electrons, corresponding to values slightly higher than the penetration velocity for the particular foil (Eq. 19·38), a large number of slow secondaries are given off on the far side of the foil, accompanied by a negligible number of primary electrons. The imaging system (Fig. 3·33) consists of a simple immersion lens of the type described by Johannson.[53] Voltages of the order of 10,000 volts are applied to the anode (B_2 and Z), focusing voltages differing only slightly from that of the foil (at zero potential) to the focusing elec-

[50] See reference 34.
[51] See reference 35.
[52] See reference 36.
[53] See reference 17.

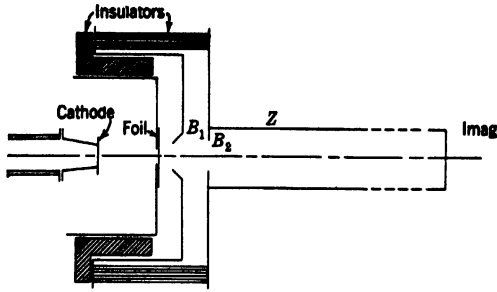


FIG. 3·33. Secondary Emission Microscope for the Study of Foils. (Behne, reference 36.)

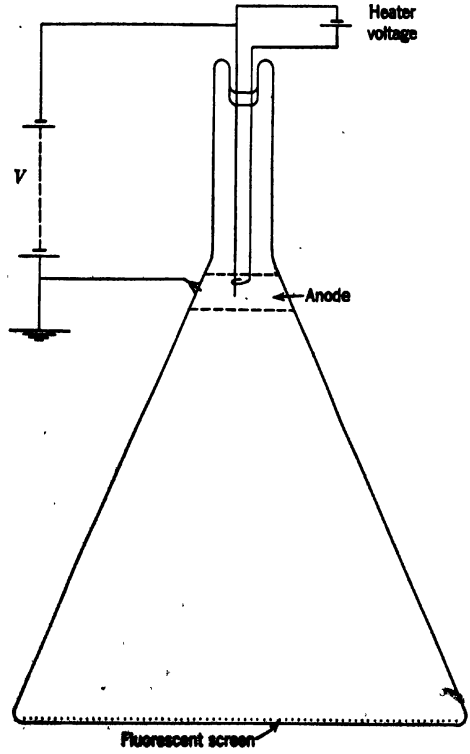


FIG. 3·34. The Field-Emission Microscope.

trode B_1. For a 0.6-micron aluminum foil the proper accelerating voltage for the primary electrons was found to lie between 5 and 7 kilovolts. Magnifications of the order of 60 were normally employed. From the sharpness of the images (about 0.5 millimeter) it was possible to conclude, making use of the formula for the chromatic aberration of cathode lenses (Eq. 16·165), that the mean kinetic energy of emission of the secondary electrons was of the order of 2 volts, in harmony with the direct measurements of other authors.

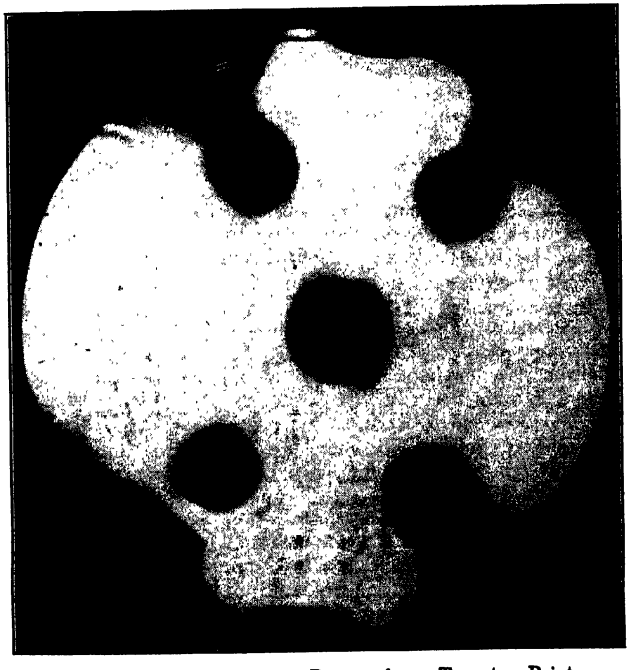


FIG. 3·35. Field-Emission Pattern for a Tungsten Point.

The field-emission microscope remains to be discussed. This instrument, devised by E. W. Müller,[54] is in essence the previously discussed point projector microscope (section 3·5) with the object removed (Fig. 3·34). The pattern appearing on the screen then shows the distribution of electron emission over the surface of the emitting point, the magnification being approximately equal to the ratio of the distance of the screen from the point to the radius of curvature of the point. Figure 3·35

[54] See reference 11.

shows a typical pattern formed by a tungsten point heated by conduction by passing a current through the supporting wires. The dark spots correspond to points of low emission (high work function) which are associated with specific crystalline orientations of the surface (the center spot corresponds to a 110-direction, the four lateral spots to 211-directions). Müller,[55] Haefer,[56] and Benjamin and Jenkins[57] have carried out extensive studies on the complex variations of the patterns with adsorption of electropositive metals as well as of gases. The patterns have quite the same character as those obtained with the later, and more complicated, apparatus of S. T. Martin[58] for studying the thermionic emission from a single-crystal sphere.

REFERENCES

1. H. Mahl, "The electrostatic electron microscope of high resolution," *Z. tech. Physik*, Vol. 20, pp. 316–317, November 1939.
2. H. Mahl, "The electrostatic electron supermicroscope," *Jahrb. der AEG-Forschung*, Vol. 7, pp. 43–56, March 1940.
3. E. Gölz, "Study of the breakdown strength of the electrode metals of the lens of the supermicroscope," *Jahrb. der AEG-Forschung*, Vol. 7, pp. 57–59, March 1940.
4. M. v. Ardenne, "An electrostatic high-voltage lens of short focal length," *Naturwissenschaften*, Vol. 27, pp. 614–615, September 1939.
5. E. Brüche, "The development of the electron supermicroscope with electrostatic lenses," *Z. Ver. deut. Ing.*, Vol. 85, pp. 221–228, March 1941.
6. C. H. Bachman and S. Ramo, "Electrostatic electron microscopy III," *J. Applied Phys.*, Vol. 14, pp. 155–160, April 1943.
7. M. v. Ardenne, "The electron scanning microscope: practical construction," *Z. tech. Physik*, Vol. 19, pp. 407–416, November 1938.
8. C. J. Young, "Equipment and methods developed for broadcast facsimile service," *RCA Review*, Vol. 2, pp. 379–395, April 1938.
9. V. K. Zworykin, J. Hillier, and R. L. Snyder, "A scanning electron microscope," *ASTM Bull.* No. 117, pp. 15–23, August 1942.
10. C. F. Eyring, S. S. Makeown, and R. A. Millikan, "Field currents from points," *Phys. Rev.*, Vol. 31, pp. 900–909, May 1928.
11. E. W. Müller, "Electron-microscope observations of field cathodes," *Z. Physik*, Vol. 106, pp. 541–550, 1937.
12. R. Haefer, "Experimental studies for testing the wave-mechanical theory of field electron emission," *Z. Physik*, Vol. 116, pp. 604–623, July 1940.
13. G. A. Morton and E. G. Ramberg, "Point projector electron microscope," *Phys. Rev.*, Vol. 56, p. 705, October 1939.
14. H. Boersch, "The shadow microscope — a new electron microscope," *Naturwissenschaften*, Vol. 27, p. 418, June 1939.

[55] See reference 11.
[56] See reference 12.
[57] See references 37 and 38.
[58] See reference 29.

15. H. Boersch, "The electron shadow microscope I: geometric-optical experiments," *Z. tech. Physik*, Vol. 20, pp. 346–350, December 1939; also *Jahrb. der AEG-Forschung*, Vol. 7, pp. 34–42, March 1940.
16. A. H. Compton and S. K. Allison, *X-rays in Theory and Experiment*, D. Van Nostrand Co., Inc., New York, 1933.
17. H. Johannson, "The immersion objective of geometrical electron optics," *Ann. Physik*, Vol. 18, pp. 385–413, October 1933; Vol. 21, pp. 274–284, November 1934.
18. W. Mecklenburg, "On the electron emission microscope," *Z. Physik*, Vol. 120, pp. 21–30, 1942.
19. M. Knoll, F. G. Houtermans, and W. Schulze, "Emission distribution at hot cathodes with the magnetic electron microscope," *Z. Physik*, Vol. 78, pp. 318–339, 1932.
20. E. Brüche and H. Johannson, "Cinematographic electron microscopy of oxide cathodes," *Ann. Physik*, Vol. 15, pp. 145–166, November 1932.
21. W. Knecht, "Combined light and electron microscope," *Ann. Physik*, Vol. 20, pp. 161–182, June 1934.
22. E. Brüche and H. Johannson, "Crystallographic investigations with the electron microscope," *Z. tech. Physik*, Vol. 14, pp. 487–488, November 1933.
23. E. Brüche and W. Knecht, "Electron-optical observation of the transformation in iron between 500° and 900° C," *Z. tech. Physik*, Vol. 15, pp. 461–463, November 1934, and "Electron-optical observation of the transformation of iron from the α- to the γ-state," *Z. tech. Physik*, Vol. 16, pp. 95–98, April 1935.
24. W. G. Burgers and J. J. A. Ploos van Amstel, "Electron-optical observation of metal surfaces":
    I. Iron: formation of the crystal pattern of activation, *Physica*, Vol. 4, pp. 5–14, 1937.
    II. Phenomena observed on transition of α into γ iron, *Physica*, Vol. 4, pp. 15–22, 1937.
    III. Crystal growth and allotropic transition in zirconium, *Physica*, Vol. 5, pp. 305–312, 1938.
    IV. Appearance of lines of high emissivity on nickel-iron crystals, *Physica*, Vol. 5, pp. 313–319, 1938.
25. J. H. McMillen and G. H. Scott, "Magnetic electron microscope," *Rev. Sci. Instruments*, Vol. 8, pp. 288–290, August 1937.
26. H. Mahl, "Electron-optical examination of electronic and ionic emission from wires," *Z. Physik*, Vol. 108, pp. 771–776, April 1938.
27. W. Heinze and S. Wagener, "Processes in the activation of oxide cathodes; I. The change in the emitting surface," *Z. tech. Physik*, Vol. 17, pp. 645–653, December 1936.
28. R. P. Johnson, "Simple electron microscopes," *J. Applied Phys.*, Vol. 9, pp. 508–516, August 1938.
29. S. T. Martin, "Thermionic and adsorption properties of tungsten single crystals," *Phys. Rev.*, Vol. 56, pp. 947–959, November 1939.
30. M. Knoll and E. Ruska, "Electron microscope," *Z. Physik*, Vol. 78, pp. 318–339, 1932.
31. E. Brüche, "Electron-optical images obtained with photoelectrons," *Z. Physik*, Vol. 86, pp. 448–450, November 1933.
32. J. Pohl, "Electron-optical images with photoelectrons," *Z. tech. Physik*, Vol. 15, pp. 579–581, December 1934.

33. H. MAHL and J. POHL, "Electron-optical images with photoelectrons," *Z. tech. Physik*, Vol. 16, pp. 219–221, August 1935.
34. V. K. ZWORYKIN, "On electron optics," *J. Franklin Inst.*, Vol. 215, pp. 535–555, May 1933.
35. E. MESCHTER, "Electron microscope for studying thermal and secondary emission," *Rev. Sci. Instruments*, Vol. 9, pp. 12–15, January 1938.
36. R. BEHNE, "Imaging of foils with the immersion objective," *Ann. Physik*, Vol. 26, pp. 385–397, July 1936.
37. M. BENJAMIN and R. O. JENKINS, "Surface migration of barium," *Phil. Mag.*, Vol. 26, pp. 1049–1062, December 1938.
38. M. BENJAMIN and R. O. JENKINS, "The distribution of autoelectronic emission from single-crystal metal points," *Proc. Roy. Soc.*, London, Vol. A176, pp. 262–279, 1940; Vol. A180, pp. 225–235, 1942.

CHAPTER 4

ELECTRON OPTICS OF HIGH MAGNIFICATION

4·1. Factors Influencing Resolving Power. One of the prime prerequisites for designing a good instrument of any type is an understanding of the factors which will limit its performance. In the case of the electron microscope the logical measure of performance is the maximum resolving power attainable.

In the present stage of development, the resolving power of the electron microscope is influenced by a number of factors. Some of these factors are functional in that they depend on the precision of construction and operation. Others are theoretical limits imposed by the characteristics of the lenses available. One factor at least, that is, the diffraction effect of the imaging pencils, is fundamental. This has already been discussed (section 3·1), and has been shown to limit the resolution obtainable with electrons as imaging medium to approximately six-tenths of a wave length of the electron radiation at the object, or, for an operating potential of 50,000 volts, to 0.03 A.U.

Allusion has also been made to the fact that the aberrations of available lens systems necessitate the use of very small imaging apertures, increasing the diffraction effects. Of these aberrations or imperfections one in particular — the aperture defect or spherical aberration of the objective — is unavoidable for lens systems of the general type employed at present. Acting by itself, it causes the spreading of an image point into a disk of confusion, whose diameter is proportional to the third power of the aperture angle of the imaging pencils. As shown in section 19·3, p. 699, the simultaneous action of diffraction and spherical aberration establishes a theoretical limit for the resolution of an electron microscope equal to about 5 A.U., the corresponding angular aperture of the objective being 0.01 radian. For the lenses employed in practice the spherical aberration is, of course, greater than that of the theoretical optimum lens. Thus the limits of resolution fixed by diffraction and spherical aberration may be higher by a factor 2 or 3 (10 or 15 A.U.) and the corresponding angular aperture lower by a similar factor (5 or $3 \cdot 10^{-3}$ radian).

Disregarding practical difficulties arising in the construction and operation and provided that the illuminating pencils have an adequately

small aperture, the attainable resolution depends on the fourth root of the spherical aberration. For this reason, as well as because this defect appears to be relatively insensitive to the shaping of the fields of the pole pieces of magnetic lenses,[1] relatively little effort has been expended in attempting to improve the performance of the microscope by reducing the spherical aberration of the lenses.

A second lens defect, which may influence the resolution and which has approximately the same value for all magnetic lenses, is the chromatic aberration. Electrons which leave the object with a velocity differing from one particular value are not focused sharply in the image plane, but form a circle of confusion about the image point, whose diameter is proportional to the difference in velocity and the effective aperture of the objective. These differences in velocity have three different origins: fluctuations in the applied voltage, variations in the initial velocities of the electrons as they leave the cathode, and energy losses experienced by them in passing through the object. For thin objects, which constitute the most favorable specimens for electron-microscope investigations, the energy losses within the object are negligible,[2] so that only the other two sources remain. Of these, the first necessitates a careful stabilization of the high-voltage equipment, and the second, the selection of a source, such as a thermionic cathode, which yields electrons of small initial velocities. Fluctuations in the coil currents of the magnetic lenses have the same effect on the image as fluctuations of the high voltage. In effect, they vary the value of the electron velocity for which the image is sharp. Thus they require a constancy corresponding to that of the high voltage.

There is one further source of local variations in the velocities of the image-forming electrons — in addition to giving rise to asymmetric deflecting forces — and thus cause unsharpness: the charging up of surfaces within the microscope in the proximity of the beam. In any imperfect vacuum, and hence in any demountable system, the passage of an electron beam is accompanied by chemical reactions in the residual gases and vapors which cause the deposition of semiconducting substances, often of a tarry consistency, on neighboring surfaces, in particular those right in the path of the beam, such as edges of apertures. Under bombardment by electrons these charge up and may exert forces on the passing electrons which are great enough to impair materially

[1] It is found that magnetic objectives with approximately symmetric fields and coil currents high enough for the pole pieces to approach saturation have, in practice, spherical aberration coefficients which differ by only a small factor from those of the most favorable calculated field distributions. See section 17·2.

[2] See section 19·2.

the quality of the image. The formation of such films around the edges of the objective aperture is a frequent source of trouble, resulting in a local deflecting field or in the formation of an irregular electrostatic lens. If this difficulty cannot be eliminated by doing without the objective aperture or by making it quite large, the objective diaphragm must be made readily exchangeable so as to permit periodic cleaning.

Other sources of image imperfections are disturbing fields of a magnetic character, such as stray fields from near-by transformers and variations in the earth's field, which may deflect the beam at right angles to the plane containing the direction of the disturbing field and that of the electron pencil considered. The elimination of this effect requires careful shielding of the electron paths and judicious placement of the microscope so as to avoid strong stray magnetic fields from electrical machines.

A very important cause of defective images which, however, cannot be classed as instrumental is found in the electrical, thermal, and mechanical instabilities of a specimen itself during exposure (section 8·1).

Finally, unless the several lenses of the microscope possess accurate axial symmetry and unless they are carefully aligned, the sharpness of the image will be impaired. This requires both precise mechanical construction and magnetic material of a high degree of homogeneity. Similarly, the mechanical construction must be such that relative displacement of different parts of the instrument — most of all, of the object support relative to the objective — will not take place. This implies both a highly rigid construction and a mounting which will minimize the effects of mechanical shock. In addition, care must be taken that elements incorporated in the rigid microscope assembly itself, such as the high-vacuum pumps, are not sources of shock.

The results of present-day electron-microscope installations show that, if all these requirements are fulfilled simultaneously, the resolutions obtained approximate the theoretically expected optimum value.

4·2. Electron Sources. It is the function of the electron source to provide a homogeneous, single beam of electrons diverging from a small area (the *effective source*) with a high current density per unit solid angle. Neither the velocity distribution in the beam nor the current density per unit solid angle[3] is affected by the subsequent refraction of the electron rays by the condenser lens. The current density per unit solid angle, in particular, determines the ultimate brightness of the image under given conditions of magnification and for a given illuminating beam aperture utilized for the image formation.

[3] The invariance of the current density per unit solid angle is a consequence of the law of Helmholtz-Lagrange, Eq. 10·30.

The beam of the electron microscope is produced by an *electron gun*, which consists of the cathode, the electron source, and accelerating and beam-forming electrodes.

Two types of electron source have found application in electron microscopes. The first of these is the cold-cathode gas-discharge tube. This electron source makes use of the secondary electrons ejected from a flat or curved metal cathode by the impact of positive ions generated in a gas discharge. The ionization, which is produced by the electrons of the beam, causes the electrons to form a narrow pencil which is accelerated toward the anode and passes out of the discharge chamber through a narrow aperture in the anode. The ions are continuously regenerated by collisions between the secondary electrons and the gas atoms, the positive carriers so formed increasing their number by subsequent collisions with other gas atoms on their way to the cathode. A diagram of such a gas-discharge tube is shown in Fig. 4·1. The cathode is usually made of aluminum, as the cathode sputtering and the consequent gradual deterioration of this electrode under ion bombardment are extremely small for this metal. Normally a gas pressure of the order of 10^{-2} mm Hg is required in the discharge chamber, this being maintained by a suitable leakage valve. The small aperture in the anode serves as a flow impedance, making possible the maintenance, in the main part of the microscope, of a much better vacuum than in the discharge chamber.

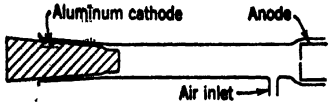


FIG. 4·1. Cold-Cathode Gas-Discharge Tube.

The source described is attractive from several points of view. No current supply is required for heating the cathode. It is insensitive to the admission of air into the instrument. The operation of the discharge is quite stable up to about 80 kilovolts. Relatively large beam currents with a small angular divergence can be attained readily. Thus current densities of the order of 1 milliampere per square millimeter for an angular divergence of 0.01 radian have been measured on a system of the type shown in Fig. 4·1.[4]

In other respects the gas-discharge tube has certain drawbacks which render it unsuitable as a source for an instrument designed for very high resolution. Unless special provisions are made, the necessary presence of a ballast resistor causes the voltage to fluctuate with the gas pressure in the discharge chamber (with the "hardness" of the tube). Furthermore, the secondary electrons emitted by the cathode have a wide range

[4] See Malsch, reference 1.

of initial velocities, corresponding to accelerating voltages as high as 10 volts.

For the above reasons the ion-bombarded cathode, in modern electron microscopes, has been replaced almost entirely by the second type of source, the thermionic cathode. Although here, again, a choice presents itself between two alternatives — the activated (oxide-coated or thoriated-tungsten) cathodes and the pure metal (tungsten or tantalum) cathodes — the conditions in the electron microscope are such that the plain tungsten cathode has been given preference, in spite of its higher operating temperatures and the correspondingly greater velocities of

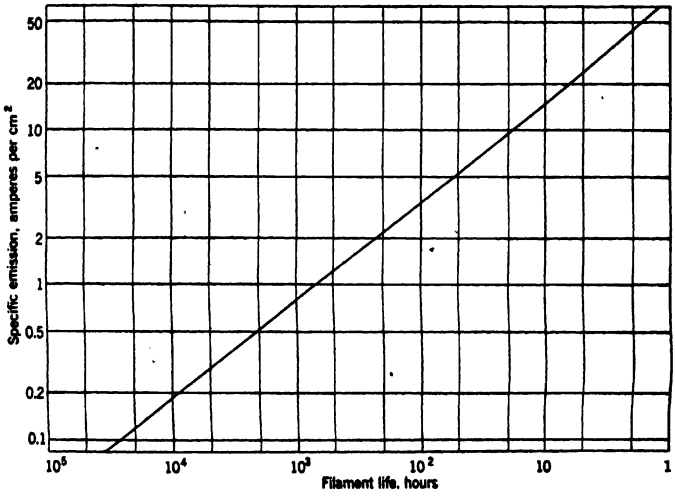


FIG. 4·2. Relation between the Specific Emission and the Mean Life of a 0.004-Inch Tungsten Filament.

emission of the thermionic electrons. The vacuum conditions in the electron microscope do not permit the maintenance of a good state of activation of an oxide-coated type of cathode, and, moreover, the contamination of parts of the microscope by fractions and products of the oxide coating would increase the difficulties from charging-up effects. A final point in favor of the pure tungsten cathode is the high specific emission attainable if the cathode is made readily interchangeable, so that a short life can be tolerated. Figure 4·2 shows the variation of the specific electron emission of the surface of a 4-mil tungsten wire with its life, which is set equal to the time required to evaporate 12 per cent of the metal of the wire. A life of 10 hours may be considered adequate in

a practical electron microscope. However, the filament life is usually considerably greater than this.

A favorable shape of the cathode is obtained if the filament is bent in the form of a " V." The vertex of this " hairpin filament " is roughly spherical. This vertex is its hottest part and forms an electron source of small dimensions, which can be centered easily in the aperture of the guard or the grid cylinder (Fig. 4·3). If the filament is heated with

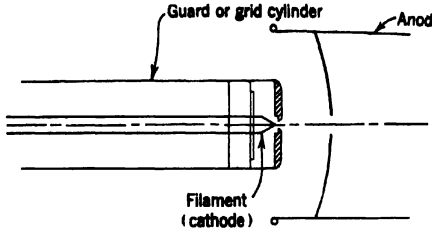


FIG. 4·3. Hot-Cathode Electron Source for the Electron Microscope.

alternating current, a symmetrical arrangement of the current supply leaves this critical point at a constant potential. The conditions are especially favorable if alternations of very high frequency are employed, as then the grid or guard cylinder provides an effective shield for the magnetic field of the heating current in the filament. Actually, the only effect of using alternating current is a slight elongation of the effective source. This is due to the deflection of the electrons as they leave the filament with low velocities by the magnetic field of the heating current.

The apertured cylinder which surrounds the filament is considered either as a guard cylinder, if it is maintained at cathode potential, or as a control grid, if its voltage relative to the cathode may be varied. The second arrangement is desirable, in particular, if no condenser lens is provided for varying the convergence of the electron beam. Beyond this cylinder, at a distance determined by the operating voltage of the instrument, the anode is located. Its aperture, which is rarely made so narrow as to limit the beam diameter, marks the entrance of the electron beam into the main body of the microscope. For practical reasons the anode and main body of the microscope are maintained at ground potential, whereas the cathode and grid or guard cylinder are connected to the negative terminal of the high-voltage supply.

The intensity and divergence of the beam leaving the anode are determined both by the position of the filament relative to the guard cylinder or grid aperture and by the potential of the latter. They are

also influenced by the space charge in the neighborhood of the cathode. It might be pointed out that this space charge has the additional effect of reducing the consequences of the asymmetries of the cathode structure, but, at the same time, lowers the effective intrinsic intensity of the source.

Obviously, an exact analysis of the beam formation in this structure would be extremely complex. It is easy, however, to consider the situation qualitatively. For a flat cathode in the plane of the grid aperture, the latter being maintained at cathode potential, the focusing fields approximate those of an apertured accelerating electrode in front of a plane at cathode potential. If the diameter of the anode aperture

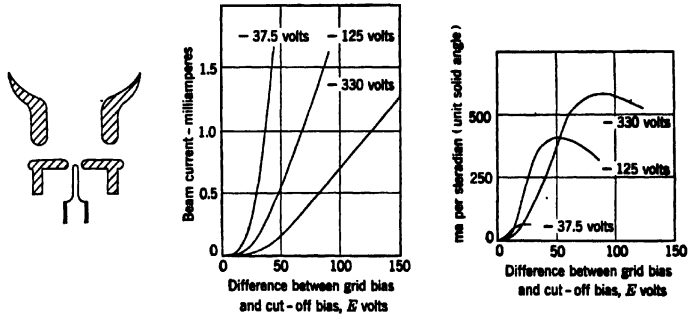


FIG. 4·4. Beam Current and Angular Concentration as Function of Cathode Position and Grid Voltage. (*a*) Beam Source. (*b*) Beam Current as Function of Grid Voltage. (*c*) Beam Current per Unit Solid Angle as Function of Grid Voltage. (Schwartz, Strübig, and Paehr, reference 2.)

is equal to its distance from the cathode, the principal rays appear to diverge from a point behind the cathode at about 2.4 times the distance between the anode aperture and the cathode (Fig. 13·24). This point represents the virtual crossover of the beam. As the cathode is drawn back into the grid shell, the divergence of the principal rays passes over into convergence. The crossover then lies in front, rather than back, of the grid aperture. At the same time the current decreases because of the reduced electrostatic field at the surface of the cathode and the correspondingly increasing effectiveness of the space-charge cloud in suppressing the emission of electrons.

The effects of changing both the grid potential and the position of the cathode relative to the grid aperture are shown in Fig. 4·4 for the system

sketched at the left. This system, representing the beam source for a television projection tube,[5] has properties fundamentally similar to those of an electron-microscope gun. The parameter in both sets of curves is the bias which must be applied to the grid to cut off the cathode emission completely, whereas the abscissas indicate the difference E between the actual grid voltage and this cut-off voltage. The first plot shows the total output current as function of E, the second the current per unit solid angle, which is the quantity of greatest importance in the electron microscope. It is seen that, although the total current increases indefinitely as the grid bias is made less negative, there is a definite optimum point for each position of the cathode (corresponding to the different cut-off voltages -37.5, -125, -330 volts) for the emission per unit solid angle. Within the range considered, this optimum value increases as the filament is moved in the direction of the anode.

The diameter of the effective electron source formed by the electron gun, that is, the crossover of the beam, is, according to Eq. 2·4, proportional to the square root of the ratio of the initial kinetic energy to the accelerating potential. For particularly simple configurations and operating conditions the coefficient of proportionality can be determined. Thus, if the cathode may be regarded as flat and the anode is a disk with a small aperture a distance l in front of it, the radius of the (virtual) crossover becomes, by Eq. 12·31*b*, $4l\left(\frac{V_r}{V}\right)^{\frac{1}{2}}$, where eV_r is the initial value of the kinetic energy of the electrons associated with motion in a lateral direction.[6] For a tungsten cathode the mean value of V_r is 3000/11,600 = 0.26 volt, 3000° A being the operating temperature of the cathode and 11,600° A the temperature for which the mean kinetic energy eV_r is just 1 electron volt. By assuming that $V = 50{,}000$ volts and $l = 1$ centimeter, it is found that the crossover diameter is about 0.2 millimeter. In practice it may be either larger or smaller than this value, depending both on the field strength at the surface of the cathode and on the strength of the focusing field in front of it. Under the actual conditions of the electron microscope the surface field at the vertex of the cathode filament is so strong that the diameter of the crossover may be only about 0.02 millimeter.

A word should be added concerning the mechanical adjustment of the electron source. It is essential that the axis of the beam should coincide exactly with the optic axis of the remaining lenses of the microscope. For this reason it is desirable to provide means both for displacing the source in a horizontal plane, normal to the optic axis,

[5] See Schwartz, Strübig, and Paehr, reference 2.

[6] Under these circumstances the crossover lies a distance $3l$ behind the cathode.

and for pivoting the system about some point near the cathode aperture. The first displacement brings the effective source on the optic axis; the second tilts the beam about the source until it is properly aligned with the axis.

4·3. The Condenser Lens. It is the function of the condenser lens to regulate the intensity and convergence of the electron pencils illuminating the individual points of the object. In addition, it may serve the purpose of shifting from bright-field to dark-field illumination. These several functions can best be described by reference to the schematic diagrams in Fig. 4·5. In each case an effective source (crossover) 0.05 millimeter in diameter is assumed to be located in a fixed position, close to the grid aperture.

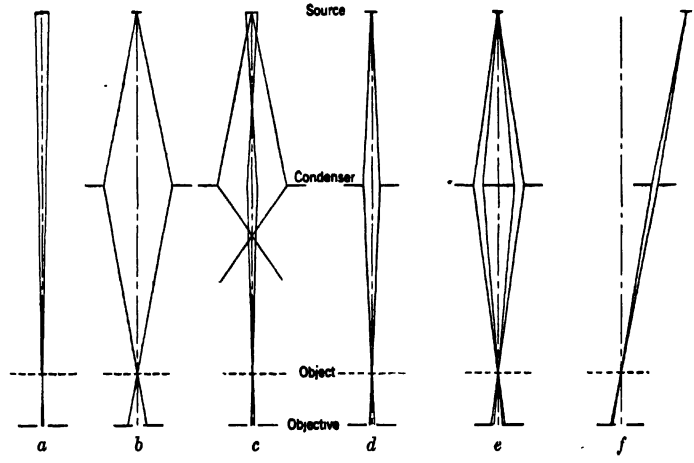


FIG. 4·5. Effect of Condenser on Illumination. (*a*) No Condenser. (*b*) Condenser Imaging Source on Object, Large Condenser Aperture. (*c*) High Condenser Current, Large Condenser Aperture. (*d*) Condenser Imaging Source on Object, Narrow Condenser Aperture. (*e*) Condenser with Ring Aperture for Dark Field. (*f*) Condenser and Source Shifted for Dark Field.

The condition *a* represents an illuminating system without condenser. The aperture angle of the illuminating pencils is seen to be given by the ratio of the crossover radius to the separation of the crossover and the object; for example, for a separation of 10 centimeters the aperture becomes $2.5 \cdot 10^{-4}$ radian. This is quite adequately small for obtaining high-resolution images. However, without condenser, it is not possible to attain a very high current concentration on the object. Even in the

most favorable case, namely, when the grid bias is adjusted to form the crossover on the object, the current density is small.[7]

If a condenser lens of large free aperture is employed to image the crossover on the object (condition *b*), the illuminated region of the object usually is of the same order as the area of the crossover. The corresponding beam divergence is large. This leads to high intensity and, although resulting in images of limited resolution, lends itself well to visual observation. In addition, the relatively large angles of the imaging pencils make the image sharpness more sensitive to changes in the focal length of the objective, so that the determination of the correct focusing currents of the objective coil is simplified. The illumination is confined to a region of the object which is not too much larger than that observed at any one time so that little beam current is wasted in heating up the object support by bombarding regions outside that actually examined. The condenser aperture should be so dimensioned that the maximum intensity is attained consistent with no appreciable impairment of the image sharpness for visual observation and not too rapid destruction of the specimen. The physical aperture for achieving this is actually quite large. For example, even if the aperture of illumination is to be only $3 \cdot 10^{-3}$ radian and if the distance between the condenser lens and the object is 16 centimeters, the condenser aperture should be 1 millimeter in diameter.

For optimum resolution in the photographic recording of an electron micrograph of the specimen after the objective has been focused, a smaller illuminating aperture is desirable. This may be attained either by exchanging the condenser aperture for a smaller one (*d*) or, more conveniently, by either increasing or decreasing the coil current of the condenser lens. If the coil current is decreased, the condition *a*, corresponding to the complete absence of the condenser, is reached in the limit, the illuminating angle being given by the cone subtended by the crossover at a point of the object. If it is increased, on the other hand, the angle of illumination is given eventually by the cone subtended by the reduced image of the crossover, formed by the condenser lens. Thus smaller convergences can be attained than without any condenser lens. Figure 4·6 shows the variation of the intensity of illumination as well as of the convergence angle of the illuminating pencil at the center of the object with a change in the focal length f_c of the condenser. The latter is related to the coil current i_c by

$$f_c = \frac{kV}{i_c^2} \qquad [4\cdot1]$$

[7] This is due mainly to the relatively large size of this crossover.

the constant k depending on the geometry of the pole pieces and on the number of turns in the coil. It is here assumed that the magnetic field intensities are too low to saturate the pole pieces, a condition normally fulfilled in the case of the condenser lens.[8] The value taken for the

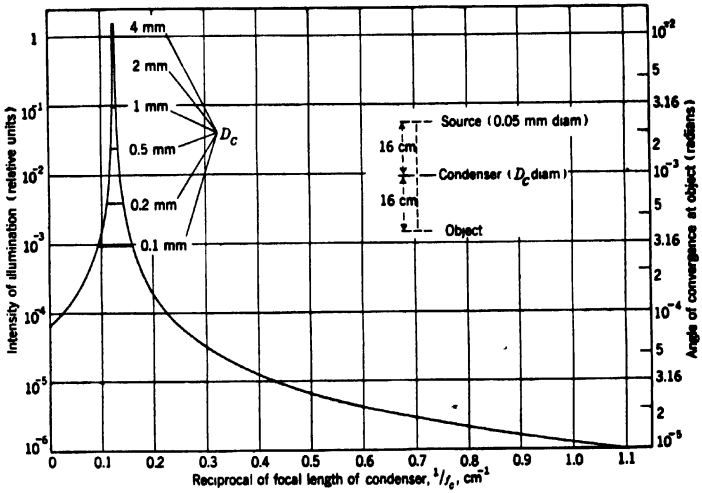


FIG. 4·6. Variation of Intensity and Aperture of Illumination with the Focal Length of the Condenser.

crossover diameter is 0.05 millimeter, for the crossover-to-condenser distance, 16 centimeters, and for the condenser-to-object distance, 16 centimeters. The value of the diameter D_c of the physical condenser aperture is indicated on the curves as a parameter. It is assumed throughout that the beam furnished by the electron gun has a uniform current distribution and a sufficient divergence to fill the condenser aperture.

The arrangements of the condenser system shown in Fig. 4·5, *e* and *f*, serve to yield a dark-field image. The parts of the image field which are free from matter appear dark, since the physical aperture of the objective is too narrow to permit the inclined direct pencils to participate in the image formation. In *e* this is accomplished by inserting an annular aperture in the condenser, the central stop being made so large that the inclination of the illuminating rays passing through the clear part of the diaphragm is greater throughout than the aperture angle

[8] In addition, the focal length of the condenser lens must be large compared to the extent of its lens field if Eq. 4·1 is to apply.

accepted by the objective. Similar results could be attained by translating a normal condenser aperture horizontally until its portion nearest the axis coincides in position with the rim of the central stop in *e*. A more intense illuminating beam can be obtained, however, by translating both the condenser lens and the source until the beam strikes the object at the required inclination (*f*).

As previously mentioned, the condenser lens differs in general from the objective and the projector lenses in having larger dimensions and simpler pole pieces. These are usually simply two cylinders separated by a nonmagnetic gap, forming part of the coil casing. This design is adequate since the focal lengths need not be so short as for the other lenses. The magnitude of the geometric aberrations as well as the constancy of the focal length of this lens are similarly of lesser importance. Here, again, the necessity for alignment of the whole illuminating system with the objective makes it desirable to provide means for horizontally displacing the condenser together with the electron gun mounted above it.

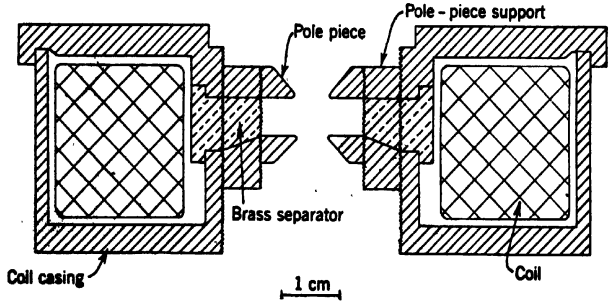


FIG. 4·7. Electron Microscope Objective (Schematic). (Ruska.)

4·4. The Objective. As in the light microscope, the objective of the electron microscope is the most important element of the whole instrument. Its properties determine, in particular, the ultimate limit of resolution which can be attained.

Like the condenser and the projector, the objective of a magnetic microscope is formed by an iron-encased coil with a suitable nonmagnetic gap. In the objective, as well as in the projector, pole pieces of highly permeable material are attached to this casing to restrict the effective gap to a very small region surrounding the electron beam. An early objective design by Ruska[9] is shown in Fig. 4·7.

[9] See reference 3.

The purpose of the casing and of the pole pieces is to increase the intensity of the magnetic field on the axis and to concentrate it on a small portion thereof, thus permitting the attainment of shorter focal lengths.[10] If iron were infinitely permeable, the focal length of the objective could be decreased indefinitely by reducing the dimensions of the pole pieces, keeping the number of turns n in the coil and the current I flowing in it — in short, the number of *ampere-turns* nI — constant. As shown in section 14·2, the magnetic field corresponds in this case exactly to the electric field between electrodes of the shape of the pole pieces, between which a potential difference $4\pi nI/10$ — the *magneto-motive force* — is applied.

The above condition is approached closely for low values of the ampere-turns nI. The focal length of the magnetic lens is, in this case, very nearly equal to that of the field formed by a single circular loop of wire carrying a current of nI amperes and having a diameter equal to the clear diameter d_p of the pole pieces.[11] This is illustrated by Fig. 4·8,

[10] Since the field integral along the axis depends only on the number of ampere-turns:

$$\int_{-\infty}^{\infty} H\,dz = \frac{4\pi nI}{10} \quad (I \text{ in amperes}) \qquad [4\cdot2]$$

the increase in field intensity is a direct consequence of the concentration. If the shape of the distribution remains unaltered the intensity must be inversely proportional, for a given number of ampere-turns, to the field length. As the focal length f is given, for " short " magnetic lenses, according to Eq. 15·86 by

$$\frac{1}{f} = \frac{0.022}{V}\int_{-\infty}^{\infty} H^2\,dz \qquad [4\cdot3]$$

the increase in the field resulting from its being concentrated into a narrow range, leaving $\int H\,dz$ constant, causes an increase in the integral for $1/f$ and a reduction in f.

[11] By Eq. 14·35 the axial field of a loop of radius $a = d/2$ is

$$H_z = \frac{2\pi nIa^2}{10(z^2 + a^2)^{3/2}}$$

and, hence, its focal length f_o is given by (provided $f_o \gg a$)

$$\frac{1}{f_o} = \frac{0.022}{V}\int H^2\,dz = \frac{0.022}{V}\cdot\frac{3\pi^3 n^2 I^2}{100d} = \frac{0.0206 n^2 I^2}{dV}$$

It is thus possible to express the focal length of the iron-encased lens by

$$f = \frac{48.4 d_p V}{n^2 I^2}\cdot k \qquad [4\cdot4]$$

where $k = d/d_p$, d being the diameter of the equivalent simple loop, applying the *short-lens* formula for the focal length. The factor k may be designated as the *form factor* of the lens.

derived by Ruska from measurements on the objective shown in Fig. 4·
The form factor k is the ratio of the focal length of the objective to th
focal length as calculated by the short-lens formula 4·3 of a lens forme
by a single loop of equal clear diameter carrying current nI. The curv
k_1 applies for very small values of nI or large values of the focal length f
The fact that the value of k_1 increases as the pole-piece diameter i
reduced may be ascribed to the simultaneous change in the shape of the
pole pieces. In any case it is seen that the form factor k does not depart
greatly from unity. The measurements were carried out with operating
voltages of 40 to 60 kilovolts.

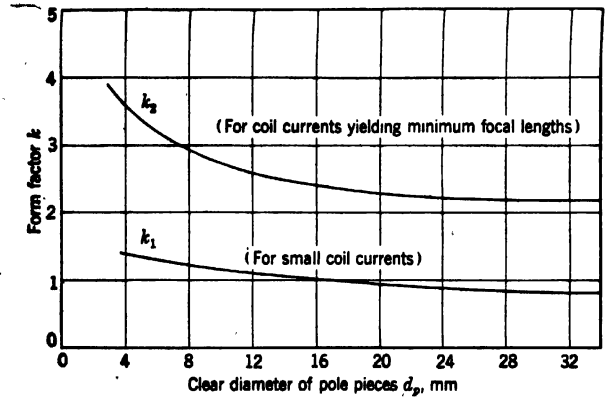


FIG. 4·8. The "Form Factor" k for Magnetic Lenses with Different Pole-Piece Diameters. (Ruska, reference 3.)

Consider now the effect of increasing the current in the coil. If the permeability of the pole pieces and the casing did not change with the value of the coil current (or the magnetomotive force), the focal length would decrease continuously, though generally at a diminishing rate,[12] as illustrated by Fig. 4·9. The abscissas are the ratio of the square of the maximum magnetic field on the axis to the accelerating potential, which, in the absence of magnetic saturation, is directly proportional to the square of the coil current, the ordinates, the values of the refractive power, $1/f$, of the lens. These values were calculated for cylindrical pole pieces separated by a narrow gap.

In practice the permeability of all known ferromagnetic materials

[12] As the focal point (object position) comes to lie in the region beyond the field maximum, the focal length continues to decrease only if the rate of decline of the field along the axis is more rapid than exponential. For an exponentially decreasing field the focal length remains constant in this region.

goes through a maximum with increasing fields as shown by Fig. 14·11. Thus, for extremely high fields, the iron becomes progressively less effective in concentrating the magnetic lines of induction, and the maximum field strength on the axis presently ceases to increase with the coil current. Instead, the field distribution along the axis broadens. Under

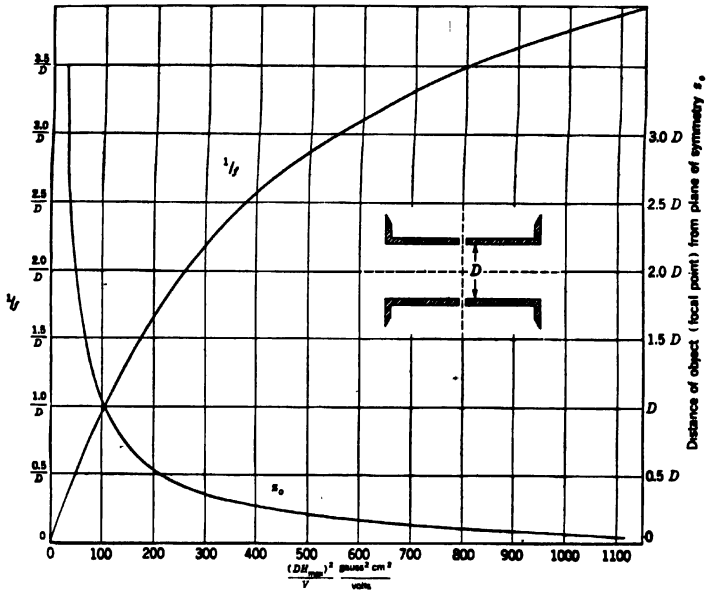


FIG. 4·9. Refractive Power $1/f$ and Object Position z_o of Simple Cylindrical Gap Lens as Function of Maximum Field Strength on Axis, Neglecting Saturation Effects.

these circumstances the focal length of a lens for electron velocities high enough to cause the focal point to fall well outside the field distribution continues to decrease, though at a progressively less rapid rate. If, on the other hand, the focal point comes to lie within the field distribution, the broadening of the field distribution will cause an actual increase in the focal length as the coil current is increased above a certain value.

Figure 4·10 shows this effect, for several voltages, for a lens with a clear pole-piece diameter of 5 millimeters and a pole-piece separation of 4 millimeters.[13] The abscissa b is the distance of the object from the center of symmetry of the lens, the distance of the image (screen) from the same point being maintained constant at 56.3 centimeters. The

[13] See Ruska, reference 3.

ordinates represent the image magnification and the coil current, respectively. The image distance being large compared to the object distance, the focal length is given very nearly by the quotient of the image

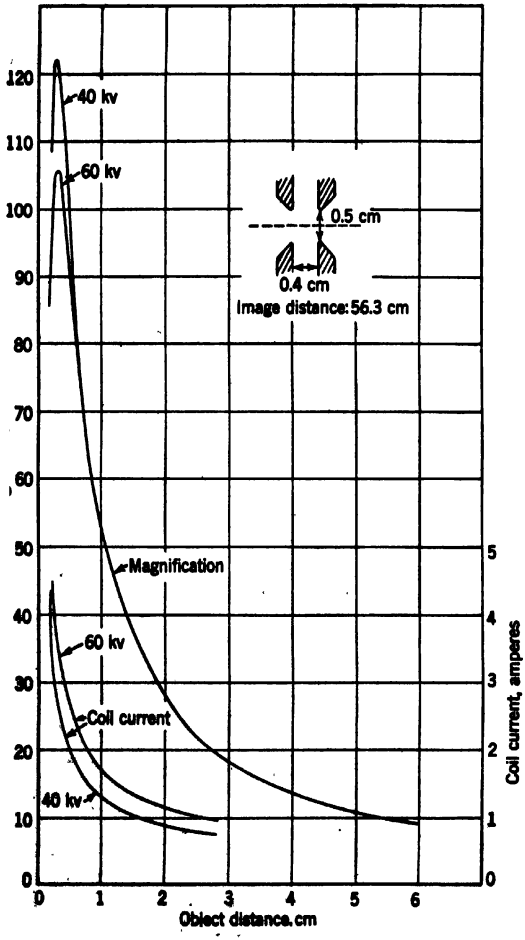


FIG. 4-10. Variation of Magnification and Coil Current with Object Distance for Fixed Image Distance. (Ruska, reference 3.)

distance 56·3 centimeters and the magnification. It is seen that for higher voltages the maximum magnification is less — the minimum focal length greater — than for lower voltages. This is a consequence of the

fact that a higher field strength is required to obtain the same focal length with the higher voltage, and the pole pieces are more fully saturated in the latter case. A plot of the minimum focal length as function of voltage, from measurements on the same lens, is shown in Fig. 4·11. The same quantity, as a function of the pole-piece diameter d_p, is plotted in Fig. 4·12. The increased effect of saturation for smaller pole pieces is shown in the sharp rise of the ratio $f_{\min}/d_p$ as d_p is reduced to small values. The curves for the minimum focal length do not pass through the origin, but appear to approach a limiting value of $f_{\min}$ equal to 2 or 3 millimeters as the pole-piece diameter is reduced indefinitely. Little weight should be attached to the value of this intercept, since it involves a somewhat uncertain extrapolation.

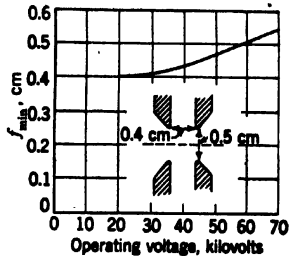


FIG. 4·11. Variation of Minimum Focal Length with Voltage. (Ruska, reference 3.)

It is clear that the same *form factor* k cannot be applied to the pole-piece lens when it is adjusted for its minimum focal length as when it is adjusted for long focal lengths. The curve k_2 in Fig. 4·8 shows Ruska's determination (at 40 and 60 kilovolts, the curve representing mean values) of the form factor for the minimum focal length. This curve, in conjunction with Eq. 4·4, indicates the value of the coil current required to obtain the minimum focal length when this focal length can be estimated.

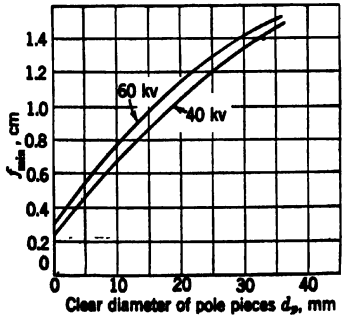


FIG. 4·12. Variation of Minimum Focal Length with Clear Pole-Piece Diameter. (Ruska, reference 3.)

Ruska[14] found a ratio of 4/3 between the pole-piece diameter and the pole-piece separation most favorable for attaining short focal lengths. For both smaller and larger separations the field distribution along the axis is flattened out to a greater degree. If the gap is large, the field strength decreases in it, and also on the axis, in inverse proportion to the separation of the pole pieces. If it is small, the leakage field of the gap contributes less and less to the axial field, so that, in the limit, the axial field approaches the distribution along the

[14] See reference 3.

axis due to the coil in the absence of iron. It should be stressed that Ruska's optimum ratio of 4/3 between the clear diameter and the pole-piece separation applies only in a very limited range.

Some recent measurements[15] have been carried out by Dosse on Siemens objective pole pieces of an early type (Fig. 4·13). His method is illustrated in Fig. 4·14. The specimen cartridge is provided with a tubular projection permitting it to explore the full length of the lens field.

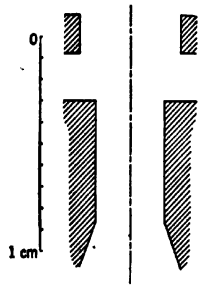


FIG. 4·13. Profile of Pole Pieces of Objective Studied by Dosse.

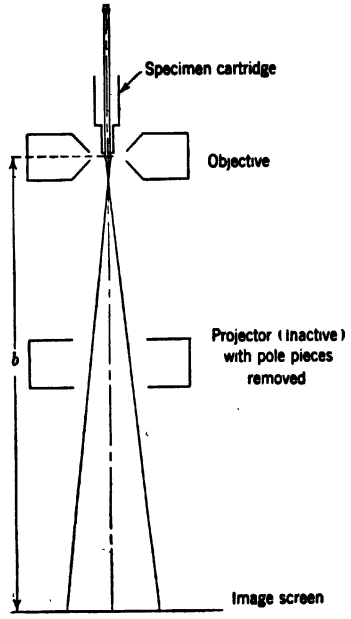


FIG. 4·14. Measurement of the Focal Length of an Objective. (Dosse.)

A disk with a fine aperture of known dimensions is mounted at the bottom of the projection and acts as object. With the projector pole pieces removed and the projector coil disconnected, the cartridge is displaced until a sharp image of the object aperture is formed on the fluorescent screen, a distance b from the center of the objective lens being tested. The magnification M of the image, which is measured, is seen to be given by $(b - f)/f$, so that the focal length f, in turn, is given by $b/(M + 1)$.

The results of Dosse's measurements are collected in Fig. 4·15. They cover the variation of the focal length with the number of ampere-turns in the lens coil for four different operating voltages ranging from 12.8 to 75 kilovolts. It is seen that, for each operating voltage, a distinct

[15] See Dosse, reference 4.

minimum is attained. In the range of voltages considered this minimum increases approximately in proportion to the square root of the operating voltage.[16]

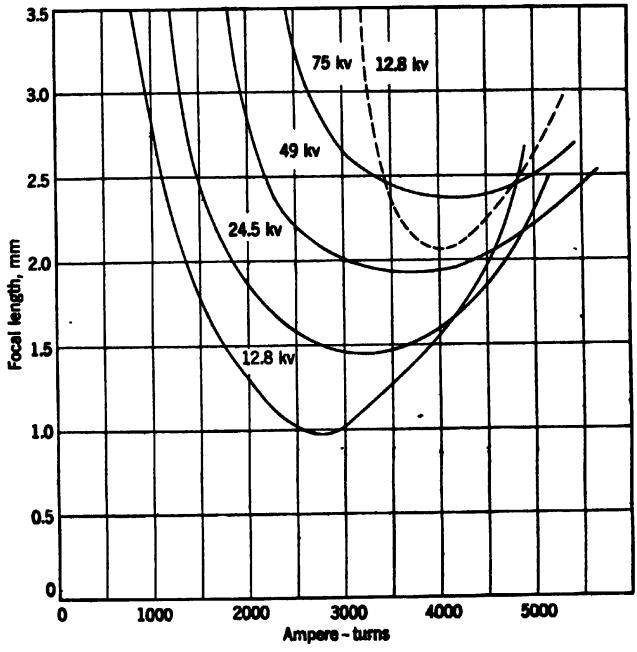


FIG. 4·15. Variation of Focal Length with Number of Ampere-Turns for Various Operating Voltages. (Dosse.)

The significance of the dotted curve in Fig. 4·15, applying for an operating voltage of 12.8 kilovolts, may be explained with the aid of Fig. 4·16. Here electron rays are represented which, in the region beyond the objective, are parallel to the axis and a distance h from it. If the object is to be imaged sharply at a great distance from the lens, it must be placed at the intersections of these rays with the axis — the focal points of the lens. Furthermore, if the angle formed by the rays

[16] Measurements by v. Ardenne (reference 5) suggest that for voltages above 60 kilovolts the minimum focal length is more nearly proportional to the voltage than to the square root of the voltage.

with the axis at the point of intersection is α_o, the focal length of the objective lens will be, by definition, $h/\tan \alpha_o$. Figure 4·16*a* represents the normal condition, in which the focal point F is located somewhat to the left of the center of the lens. As the number of ampere-turns in the lens coil is increased, the focal point moves to the right. In Fig. 4·16*b* it lies just at the center, so that the ray is parallel to the axis on both sides of the lens. For still greater exciting currents (Fig. 4·16*c*) the ray intersects the axis at two points, F_1 and F_2; for these object positions

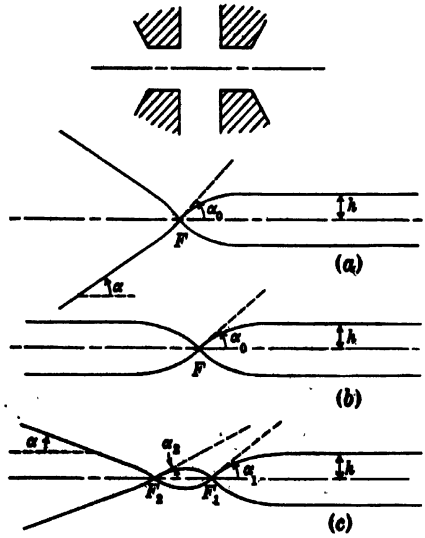


FIG. 4·16. Ray Paths in the Field of a Magnetic Objective (Schematic).

the focal lengths of the lens become $h/\tan \alpha_1$ and $h/\tan \alpha_2$, respectively. The full-line curves in Fig. 4·15 relate to the object position F or F_1, whereas the dotted-line curve refers to the object position F_2. The first focal length reaches a minimum for an exciting current close to that leading to the condition shown in Fig. 4·16*b*. Beyond this point it increases again, whereas the second focal length decreases from infinity down to a minimum which, owing to the flattening out of the field distribution curve, lies considerably higher than the minimum of the first focal length. Additional focal points F_3, F_4, $\cdots$ and corresponding focal lengths arise as the current in the lens coil is increased and the intersections of the imaging rays with the axis multiply. The normal

operating range is given by the declining portion of the full-line curves in Fig. 4·15. The region very near the minimum is unfavorable since here the portion of the objective lens field which lies between the object and the source causes the illuminating rays to converge on the specimen so as to make the aperture of illumination excessively large.

The principle on which the pole pieces and the casing of the coil are designed is that the *magnetic resistance* offered to the flux by the iron part of the circuit is small compared with that offered by the gap.[17] At the same time, the cross-section area of the gap must be made large enough for the field on the axis in the center of the gap to attain the largest possible value. This has led to pole-piece designs of the general character shown in Fig. 4·17.[18] A relatively narrow channel through the pole pieces permits the passage of the electron beam. The pole-piece tips are ground off flat to a diameter considerably greater than that of their free opening. In this manner saturation effects in the pole-piece tips are reduced to a minimum. In view of the large diameter of the casing and the lower value of the magnetic induction within it (resulting in a higher value of the permeability) its thickness need not be made very large, 5 millimeters being adequate for practically all purposes. If emphasis is to be placed on reduction in weight, disregarding all practical difficulties of production, the ideal materials for the casing and pole-piece base would be an alloy of extremely high permeability at moderate field strengths, such as Permalloy in Fig. 14·11. The pole-piece tips, on the other hand, should be made of a substance maintaining high per-

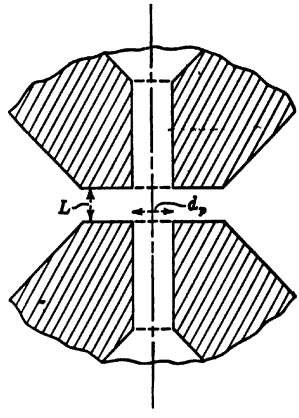


FIG. 4·17. Typical Pole-Piece Assembly of a Modern Objective or Projector.

[17] Mathematically this may be expressed by

$$\int_{\text{iron}} \frac{dx}{\mu S_m} \ll \int_{\text{gap}} \frac{dx}{S_g}$$

where dx is an element of length of the magnetic circuit, μ the permeability, S_m the cross-section area of the iron at any point, S_g the effective cross-section area of the gap. S_m is here equal at any point to the product of the thickness of the iron and $2\pi R$, R being the radial distance from the axis. The expression neglects leakage flux.

[18] Recent measurements have indicated that while the above pole pieces are quite satisfactory for most practical purposes, they do not necessarily represent the absolute optimum.

meability even at large magnetic inductions, as represented by Permendur in the same figure.

Figure 4·12 indicates the variation of the minimum focal length with the free pole-piece diameter for a particular lens type. It is found that, quite generally, the minimum focal length becomes smaller as the pole-piece opening is reduced. The least focal lengths attained with objectives of modern design at 50 kilovolts are of the order of 1 millimeter, the corresponding pole-piece openings having diameters slightly smaller than this. To reach very much lower focal lengths with smaller pole-piece openings, it would be necessary to employ magnetic materials capable of greater intensity of magnetization. If the diameters of the two pole-piece openings differ, as may be suggested by the necessity of accommodating the object cartridge in one of them, the larger of the two must be expected to play the more important part in determining the maximum field concentration. No great reduction in focal length can be expected from decreasing the size of just one of the pole-piece openings.[19] Thus, a certain limitation in the focal lengths obtainable with the magnetic objective lens is imposed by the necessity of introducing the specimen into the field. Since the object mount cannot conveniently be reduced in size indefinitely, one of the pole pieces must either be given a sufficiently large clear diameter to permit the object to pass through it, or the possibility of placing the object at the point of the field corresponding to the minimum focal length must be sacrificed. Either procedure involves a restriction on the available refractive power over and above that imposed by the magnetic characteristics of the pole-piece material alone. For this reason the least focal length obtained with normal objectives is commonly two or three times as large as that reached with the projector, for which a similar limitation does not exist.[20] The minimum objective focal length, generally, lies in the range between 2 and 5 millimeters. The shortest objective focal length obtained so far, at 60–70 kilovolts, is 0.9 millimeter.[21] This requires, however, object holders of a special kind, which are considerably more difficult to handle than normal object holders.

The question of lens coil design also merits some attention. The problem consists in determining the shape of the coil cross section which will

[19] Unpublished measurements by A. C. Schroeder (RCA Laboratories) indicate that the arithmetic mean of the two diameters may be regarded as the "effective diameter" of the system. The use of very small openings in one of the pole pieces is undesirable, as it may be expected to increase the spherical aberration of the objective.

[20] In the absence of restrictions on the object position identical lenses will yield a shorter minimum focal length if used as objectives than if used as projector lenses; see section 4-5.

[21] See von Ardenne, reference 6.

lead to the maximum magnetomotive force (or maximum number of ampere-turns nI) and the least dissipation of energy. Let ρ be the resistivity in ohm-centimeters of the coil-wire material (normally copper), R_m the mean radius of the coil in centimeters, A the cross-section area of the coil in square centimeters, and V the voltage applied between its terminals. Then simple geometric considerations lead to the following formula for the number of ampere-turns:

$$nI = \frac{VA}{2\pi n R_m \rho} \qquad [4\cdot5]$$

The power dissipation, similarly, becomes

$$VI = \frac{V^2 A}{2\pi n^2 R_m \rho} = (nI)^2 \cdot \frac{2\pi R_m \rho}{A} \qquad [4\cdot6]$$

The last equation shows that the coil should be made as long and of as small inner diameter as possible. Once the outer diameter of the coil is large compared to the inner diameter, a further increase of the former does not alter the ratio R_m/A appreciably. On the other hand, the length of the coil enters the expression only through A, which is proportional to it; accordingly, the heat dissipation for a given magnetomotive force is inversely proportional to the coil length.

Normally, the required number of ampere-turns nI, the resistivity ρ, the voltage of the available supply, and the inner diameter of the coil are prescribed, and the values of n, A, and R_m must be determined to yield the permitted power dissipation. Then Eq. 4·6 yields the ratio

$$\frac{A}{R_m} = \frac{2h(R_o - R_i)}{R_o + R_i}$$

where h is the length and R_o and R_i are the inner and outer diameters of the coil, respectively. For any permissible outer diameter R_o both the length h and the cross-section area A are thus given. Equation 4·5, on the other hand, leads to an expression for the resistance per unit length σ of the coil windings:

$$\sigma = \frac{\rho n}{A} = \frac{V}{2\pi n I R_m} \qquad [4\cdot5a]$$

The number of turns n also follows from this last equation, after A has been established. The proper value of ρ is obtained by multiplying the resistivity of copper ($1.72\cdot10^{-6}$ ohm-centimeter) by the ratio of the total coil volume to that actually occupied by the copper of the windings. If the windings are staggered, the ratio of the total volume to that occu-

pied by the windings becomes $2\sqrt{3}/\pi = 1.10$. If an additional 10 per cent is allowed for the space occupied by the insulation, including the paper separators between windings, a suitable value for ρ becomes $2.1 \cdot 10^{-6}$ ohm-centimeter.

Assume, for example, that a field of 10,000 gauss is required in a gap 0.5 centimeter in length. If the magnetic resistance of the pole pieces and the casing as well as inhomogeneities of the field at the gap are neglected, a magnetomotive force of $4\pi nI/10 = 5000$ gauss-centimeters or 4000 ampere-turns is required. If, now, $V = 90$ volts, $A = 50$ square centimeters, $R_m = 5$ centimeters, and $\rho = 2.1 \cdot 10^{-6}$ ohm-centimeter, Eq. 4·5 shows that 17,100 turns with a resistance per unit length of 0.000718 ohm per centimeter are needed. If B and S No. 23 wire, which has slightly smaller resistance per unit length and slightly greater diameter, is employed, the number of turns must be reduced to 15,900 to be accommodated in the same space. The total resistance of the coil will now be $2\pi\sigma nR_m = 353$ ohms, the coil current 0.255 ampere, so that the total number of ampere-turns nI becomes 4050 and the energy dissipation in the coil, $VI = 23$ watts.

So far no mention has been made of the aberrations of the objective. Measurements and calculations indicate that both spherical aberration (section 17.2) and chromatic aberration (section 17.6) — the only two lens defects which affect the sharpness of the central portion of the image and which are important for an electron-microscope objective — are very nearly identical for the different magnetic lenses employed at the present time, provided that the focal length and the clear diameter of the pole pieces are the same. If, for a given pole-piece assembly, the focal length is reduced, both the spherical and the chromatic aberrations decrease. The reduction in the spherical aberration, in particular, leads to a decrease in the least resolvable separation almost in proportion to the focal length (Fig. 17·7). Thus, if the magnification of the image is increased by increasing the coil current of the objective and displacing the object correspondingly, the sharpness of the image is only slightly decreased. Decreasing the pole-piece assembly to scale so as to reduce the focal length leads (if spherical aberration and diffraction are the factors limiting the resolution) to an improvement in resolution proportional to the fourth root of the scale factor if the pole pieces are unsaturated. In the region of saturation the resolution must, eventually, be expected to decline with decreasing pole-piece diameter, since the reduction in focal length lags behind the decrease in the clear aperture to an increasing degree. For a given pole-piece design the optimum resolution should be obtainable with very high coil currents, of the order of that giving the minimum focal length.

4·5. The Projector. It is the function of the projector to project on the screen or photographic plate a magnified image of the intermediate image formed by the objective. The general character of the lenses employed for this purpose is the same as that of the objective discussed in the preceding section. Its design is simplified slightly by the fact that no provision has to be made to accommodate the specimen holder.

From this similarity of construction it follows that the aberrations of the projector are similar to those of the objective. Yet their effect on

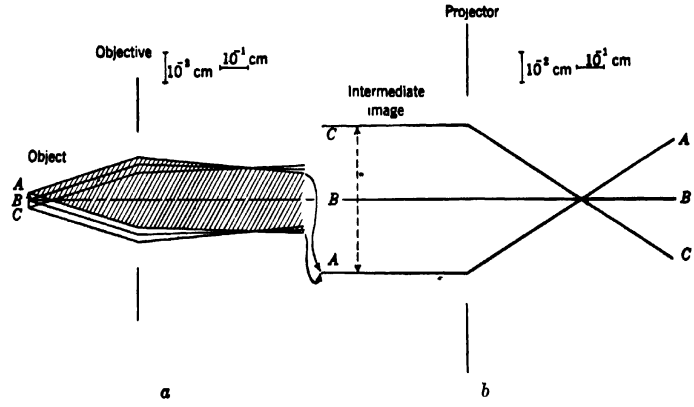


FIG. 4·18. Imaging Pencils in the Objective (*a*) and the Projector (*b*).

the image is entirely different. The reason for this becomes clear from a comparison of the ray paths in the objective and in the projector (Fig. 4·18). Distances and angles correspond to the case of two 4-millimeter focal-length lenses separated by 40 centimeters and forming a final image 5 centimeters in diameter with a total magnification of 10,000, each lens contributing the same factor 100 to the total. To represent the rays more clearly all radial distances have been multiplied by a factor 100 and 10, respectively, as compared to the axial distances. Only the portion of the object and the intermediate image, as well as the ray pencils, contributing to the formation of the final image are shown. It is seen that at the objective relatively wide-angle pencils image an object area small compared to the lens area utilized by the pencils, while at the projector very narrow pencils image an area of the intermediate image large compared to the lens area utilized by any one pencil. Thus the image defects resulting from the imperfections of the

objective are substantially the same for all the object points imaged and equal, in particular, to those for the object point on the axis. The only defects which affect the imaging of this point are, as has already been mentioned, spherical aberration and chromatic aberration. These two produce an equal unsharpness throughout the, otherwise faithful, portion of the intermediate image utilized for the final image.

At the projector, on the other hand, the aperture of the individual pencils is so small that the central point of the intermediate image is imaged, to all intents and purposes, without aberration. However, since the pencils imaging the marginal portions pass through the outer zones of the lens, they are affected by the lens defects. Whereas the aberrations of the objective determine, primarily, the sharpness of the center of the image, those of the projector may cause differences in sharpness between the axial and marginal portions thereof, as well as deviations from faithfulness in the reproduction of the image. Thus the outer pencils are brought to focus more rapidly (curvature of field) and asymmetrically (astigmatism), reducing the sharpness of the image.[22] This effect, augmented by the corresponding aberrations of the objective lens, is generally too small to be significant.

In addition, the principal rays of each pencil intersect the axis ahead of the paraxial focus of the lens, resulting in pincushion-shaped distortion.[23] This second effect (supplemented by the rotational distortion characteristic of magnetic lenses) is the more important. If the projector can be treated as a thin lens it becomes a direct consequence of its spherical aberration. If the distortion is measured by the ratio of the magnifications of the marginal portion and of the central portion of the image, it remains approximately constant (increasing slightly) as the projector magnification is increased, provided that the area of the intermediate image which is reproduced is held constant. However, if the area of the final image is kept the same, the visible distortion increases rapidly as the projector magnification is reduced.

It is hence advisable, if the observed images are to exhibit no more than a prescribed amount of distortion, to provide an aperture above the projector, which limits the field. This aperture is usually cut into a fluorescent screen on which the intermediate image can be viewed. The proper value for the diameter of this aperture can readily be estimated if the projector is treated as a thin lens. If S_3 ($= MC/f^2$) is the spherical aberration coefficient of the lens and M is the magnification on the axis, the increase in magnification for a point of the intermediate

[22] See section 17-4.

[23] See section 17-5; although the distortion is normally of the pincushion type, it is of necessity so only if the projector is a thin lens.

image a distance r from the axis is given by S_8r^2,[24] assuming that the magnification M is large compared with unity. Hence the requirement that the increase in magnification in passing from the center to the edge of the image should not exceed 10 per cent leads to a field aperture radius $r = (0.1M/S_8)^{1/2}$. With $S_8/M = 3$ diameter^{-2} (Fig. 17·5), $r =$ 0.18 clear pole-piece diameter. Thus the diameter of the field aperture should be about one-third of the clear diameter of the projector pole pieces.

The effect of chromatic aberration, that is, of variations in the kinetic energies of the electrons forming the image, is to vary the magnification in inverse proportion to this kinetic energy. This causes a linearly increasing unsharpness of the final image from the center outward, to which projector and objective contribute equally. If the kinetic energy varies by 0.01 per cent, the unsharpness, due to this cause, at the margin of a 5-centimeter image becomes only 5 microns.[25] An additional effect of about the same magnitude arises from the variation in the rotation of the image with the kinetic energy of the imaging electrons. This unsharpness is less than half the axial chromatic defect due to the objective (assuming a magnification of 10,000, an aperture of $3\cdot10^{-3}$ radian, and a focal length of 0.4 centimeter) and can thus readily be neglected.

A further interesting difference in the operation of the projector and the objective consists in the fact that the magnifying effect of the projector is determined by the complete lens field, whereas that of the objective depends only on the portion of the field beyond the object. The field "overlap" influences, in the case of the objective, only the conditions of illumination of the specimen. In the extreme case when

[24] Under the assumed conditions the ratio of the separation r' from the axis of a particular point in the image to the separation r of the conjugate point in the object (intermediate image) is given by

$$\frac{r'}{r} = \frac{L - f + \dfrac{S_8r^2f}{M}}{f - \dfrac{S_8r^2f}{M}} \cong \left(1 + \frac{S_8r^2}{M}\right)\frac{L-f}{f} = M\left(1 + \frac{S_8r^2}{M}\right) \qquad [4\cdot7]$$

Strictly speaking, the radial magnification at the margin is not given by r'/r, but by dr'/dr, for which the deviation from M is $3S_8r^2$. Thus the increase in the *overall magnification* of 10 per cent, discussed above, corresponds to an increase of 30 per cent in the local radial magnification at the margin. It should be noted, furthermore, that, if the lens is thick and the object is near its center, the coefficient of distortion may deviate greatly from the value given. Here, in addition, the anisotropic distortion, which vanishes for the very short lens, is likely to play an important role (see section 17·5).

[25] This figure, again, assumes thin-lens conditions. For the thick lenses normally employed the unsharpness is, generally, less. See section 17·6.

the intermediate image falls in the center of a symmetrical projector lens (Fig. 4·16*b*), that is, if the image-side and object-side focal points coincide, the projector lens has no magnifying action whatever, although the same lens used as an objective with the specimen placed at its center might give a very large magnification.[26] Thus the projector magnification ceases to increase with increasing number of ampere-turns in the

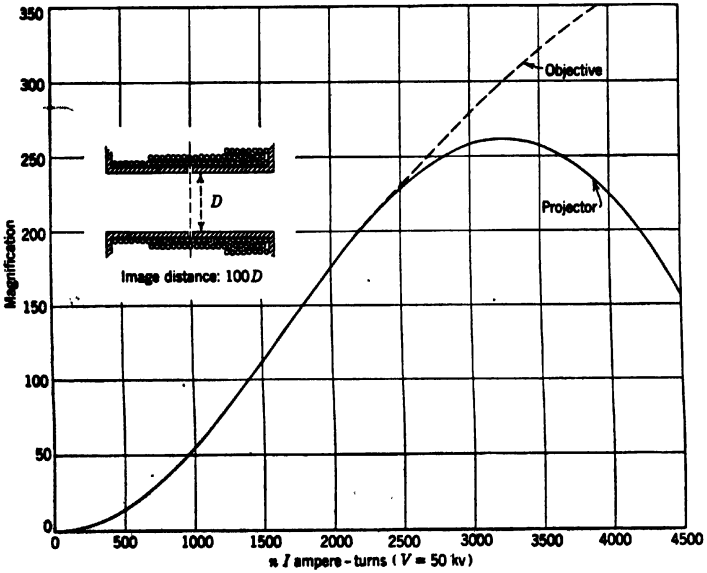


FIG. 4·19. Variation of Magnification with Coil Current for a Nonsaturating Magnetic Lens Used as Projector or Objective.

lens coil even in the absence of pole-piece saturation. This is illustrated by Fig. 4·19. Here the variation of the magnification with the number of ampere-turns in the lens coil has been plotted for a lens formed by coaxial cylindrical pole pieces of equal diameter and infinite permeability separated by a narrow gap (Fig. 17·3, System *C*) used either as an objective or as a projector. The image distance has been assumed to be 100 lens diameters in either case. Although the objective magnification is still rising at 4400 ampere-turns, the projector magnification does not increase appreciably beyond 3000 ampere-turns. Here

[26] Under these circumstances the field aberrations of the objective would be extremely large, so that, at least with the normal system of illumination, the adjustment would be unfavorable for operation.

the focal length of the lens is approximately 0.35 diameter and the separation of the focal point from the center of the lens, 0.23 diameter.

The focal properties of electron lenses employed as projectors can be examined conveniently with the aid of the system shown in Fig. 4·20. The lens pole pieces to be tested are inserted into the objective casing of an electron microscope from which the projector unit has been removed. A fine-mesh screen of known mesh constant is placed above the lens, well outside the lens field, and is illuminated with a parallel beam of electrons. The shadow projection of the wire screen formed by the electrons on the fluorescent screen or photographic plate a distance b below the lens is measured out so as to yield the magnification factor M. The geometry of the figure indicates that $M = (b - f)/f$, so that, in turn, $f = b/(M + 1)$.

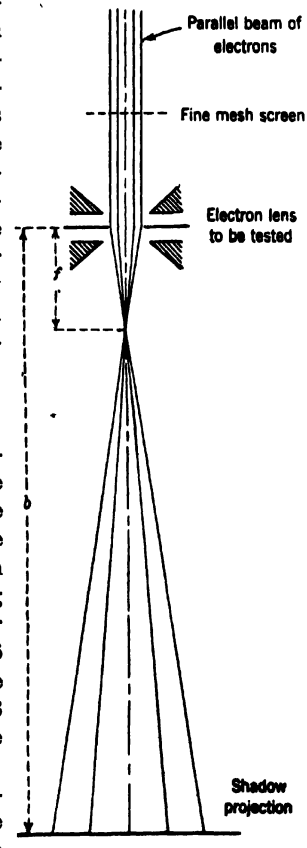


FIG. 4·20. Measurement of the Focal Length of a Projector.

This method of measurement has recently been applied in the study of the magnetic lens formed by the pole-piece system shown in Fig. 4·21. The shape of the lower pole piece was obtained in a search of a pole-piece design yielding the smallest minimum focal length for a fixed pole-piece separation of 0.236 centimeter, the clear diameter of the opening of the pole pieces being 0.318 centimeter. Both pole pieces were made of Armco iron.

Figure 4·22 shows the results of measurements of the magnification of the magnetic lens formed by these pole pieces, varying both the spacing and the coil current. The operating voltage remained fixed at 47.5 kilovolts. It is seen that for moderate coil currents the pole-piece spacing L leading to the greatest magnification (shortest focal length) is about one-third of the clear diameter. At high coil currents the magnification curve has secondary maxima at larger pole-piece spacings. Thus, for a coil current of 180 milliamperes the most pronounced

secondary maximum is located at $L = 0.236$ centimeter. It should be noted that the greatest magnification attained in this series of measurements corresponds to a focal length of 1.67 millimeters, which is only about half as great as the clear diameter of the pole-piece openings.

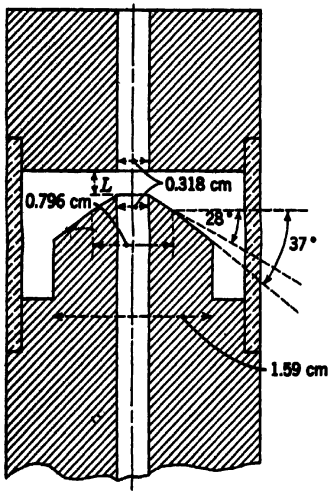


FIG. 4·21. Test Pole Pieces of Projector Lens.

The more complex character of the curves for the higher coil currents must be attributed to magnetic saturation in the pole-piece tips. It was found that the magnetic-field distribution along the axis, for a particular coil current and pole-piece spacing, was fitted, within experimental error, by the formula

$$H = \frac{H_0}{1 + \left(\frac{z}{a}\right)^2} \qquad [4\cdot8]$$

The focal length of this field distribution is directly proportional to the half-value width $2a$ if both the number of ampere-turns and the operating voltage are kept fixed.[27] Hence, if it is assumed that Eq. 4·8 represents the axial field distribution throughout, the magnification curves in Fig. 4·22 can readily be translated into curves for the half-value width. This has been done in Fig. 4·23. It is seen that for coil currents up to 120 milliamperes the half-value width increases only slightly with the exciting current. Here pole-piece saturation plays a minor role. For higher coil currents the width of the distribution increases more rapidly and, furthermore, becomes quite sensitive to variations in pole-piece spacing. The last effect may be ascribed to a variation in the distribution of the magnetization of the pole-piece tips.

The manner in which the pole-piece shape here investigated was obtained suggests that the maximum of the refractive power located at $L = 0.236$ centimeter is largely dependent on the exact shaping of the mantel surfaces of the pole pieces and the choice of the magnetic material. This is confirmed by measurements made on pole pieces with tips of Permendur (2 per cent vanadium, 49 per cent cobalt, 49 per cent iron), both heat-treated and untreated. The principal maximum of the

[27] See Eq. 15.75 with k = const.; k, for constant voltage, is proportional to aH_0 which is proportional to the number of ampere-turns.

magnification located at approximately one-third of the clear diameter is obtained in all cases. At the same time, the shape of the magnification curves and the position of the secondary maxima vary greatly. The principal conclusion of this series of studies, which has embraced

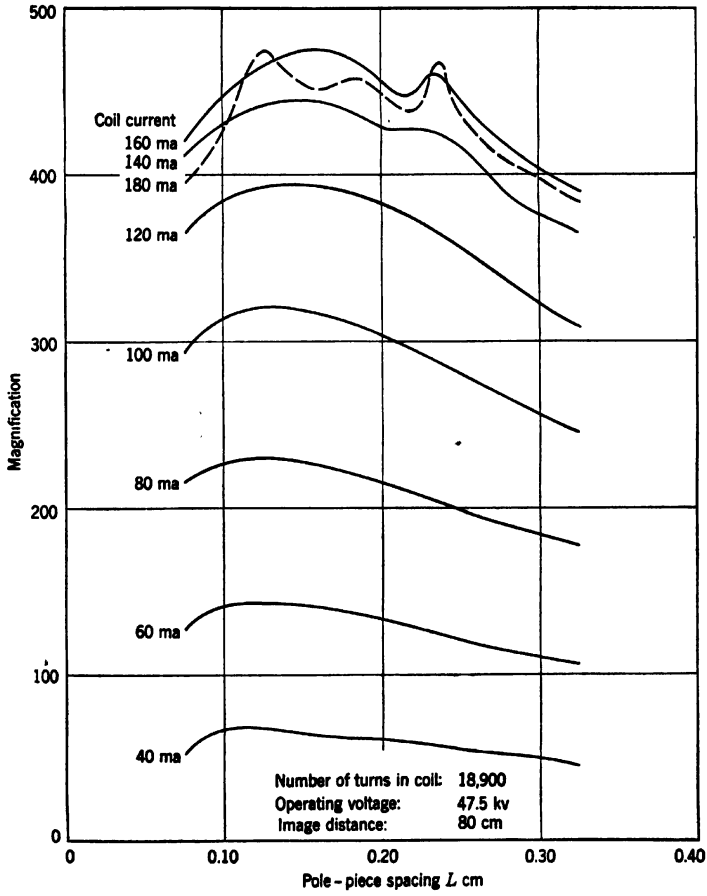


FIG. 4·22. Variation of Magnification of Magnetic Lens Shown in Fig. 4·21 with Pole-Piece Spacing for Various Coil Currents.

pole pieces of different shape as well as of different composition, is that the effective half-value width of the field distribution is critically dependent on the diameter of the pole-piece opening and practically independent of the spacing.

If a large image field at low magnification is desired, a double projector lens may be provided, as in Mahl's electrostatic microscope (section 3·3). The preferable condition is here, also, to place the high-magnification (small pole-piece diameter) lens approximately at the focal point of the low-magnification (large pole-piece diameter) lens

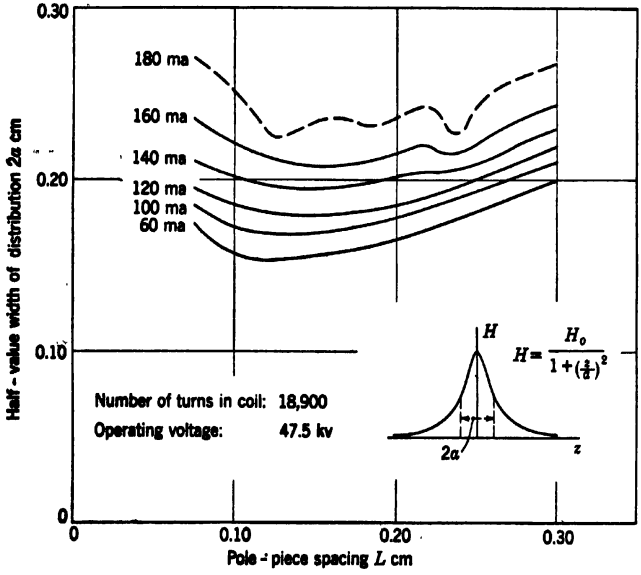


FIG. 4·23. Variation of Effective Half-Value Width of Field Distribution with Pole-Piece Spacing for the Magnetic Lens Shown in Fig. 4·21.

above it, so that the image field of the latter is not restricted by the former. This limits, it is true, the useful magnifications of the weak lens to a relatively narrow range. The flexibility of the high-magnification lens is unaffected. A shift from one lens to the other may require slight refocusing of the objective if the attainment of high resolutions is desired.

More than two lenses, with intermediate images intervening, may be required if very high magnification is to be combined with small overall length of the microscope. Optically, the lenses between the first and the last partake of the properties of both the objective and the projector. Even here, however, the objective will generally determine the ultimate limits of performance of the instrument.

4·6. Depth of Field. The very small angular aperture of the imaging pencils in the electron microscope results in a great depth of field. A sharp image is formed not only of the object structure lying exactly in the object plane, but also of that which lies some distance in front of and behind it. Let α (Fig. 4·24) be the aperture of the image-forming pencils leaving any one object point. It is seen that the pencils leaving a point a distance $D/2$ behind or in front of the object plane — that is, the plane optically conjugate to the plane containing the fluorescent screen or photographic plate — intersect the object plane in a disk of radius d, where $d = (D/2)\tan\alpha$. If, now, d is set equal to the maximum reduction in resolution due to imperfect focusing which can be permitted, D becomes the depth of field or the axial extent of the object giving a sharp image:

$$D = \frac{2d}{\tan\alpha} \qquad [4\cdot9]$$

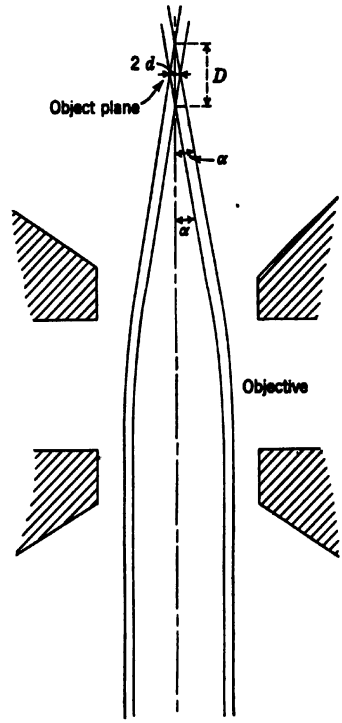


FIG. 4·24. The Depth of Field of an Electron-Microscope Objective.

For $\alpha = 3\cdot10^{-3}$ radian and $d =$ 30 A.U., $D = 2\cdot10^{-4}$ centimeter (2 microns). Thus, with a magnification of 100,000 and an image diameter of 20 centimeters, the depth range of an object imaged sharply throughout may be equal to the diameter of the portion imaged. This stands in marked contrast to conditions in the high-magnification light microscope, where, with α having values as large as 70 degrees, the depth of field may actually be less than the least resolvable separation.

Indirectly, the great depth of field reduces the stringency of the requirements on the stability of the voltage and the lens currents as well as the precision with which these quantities must be adjusted in focusing the image. Furthermore, it has the fortunate consequence that the quality of focus is, normally, unaffected by the current setting

of the projector. The lens current of the projector may be varied arbitrarily to obtain a desired magnification without changing the sharpness of the image. Assume (Fig. 4·25) that the image has been focused sharply for the maximum magnification and that, subsequently, without altering either the current in the objective coil or the position of the object, the current in the projector is reduced so as to give a lower

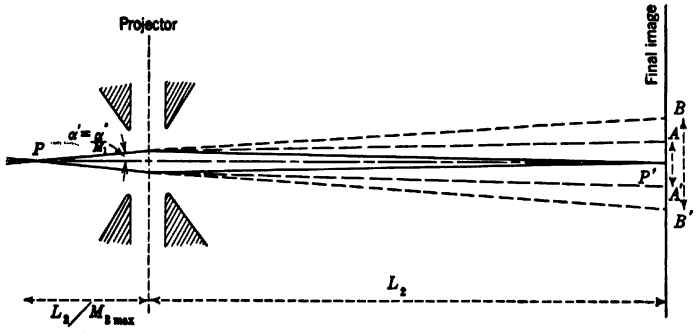


FIG. 4·25. Independence of Focusing of the Projector Coil Current.

magnification. Then the beam from the point of the intermediate image P, originally converging in P', intersects the plane of the final image in the circle of confusion AA'. Since the projector lens is in all cases positive, this is necessarily smaller than the circle BB' formed if no projector lens is present. If the aperture angle at the intermediate image is $\alpha' = \alpha/M_1$ (M_1 = magnification of objective) and the distance between the projector and the final image is L_2,

$$BB' = 2L_2\left(1 + \frac{1}{M_{2\max}}\right)\alpha' \qquad [4\cdot10]$$

where $M_{2\max}$ is the maximum magnification of the projector. Thus, for $L_2 = 40$ centimeters, $\alpha = 3\cdot10^{-3}$ radian, $M_1 = 100$, and $M_2 \gg 1$,

$$BB' \cong 2\cdot10^{-3} \text{ centimeter}$$

This is of the same order as the resolution of fine-grained emulsions. It should be noted that for low objective magnifications the dependence of the image sharpness on the projector magnification may become appreciable.

4·7. Object Thickness and Resolution. In the initial section of this chapter it was mentioned that for the very thin specimens which provide the most favorable objects for electron microscopy, the changes in velocity experienced by the electrons in passing through the object

are generally of negligible effect. This is not the case for specimen thicknesses in the range of 1000 A.U. to 1 micron,[28] where the limit of resolution, as determined by chromatic aberration alone, varies from 50 A.U. up to the limit of resolution of the light microscope (about 1000 A.U.) and, for dense materials, even beyond this value. If no physical limiting aperture is employed in the objective, the broadening of the angular distribution of the imaging pencils will give rise to a further loss of sharpness due to both spherical and chromatic aberration.[29]

The volume scattering of the electrons within the object is a further source of unsharpness which depends on the thickness of the object. Viewed from the objective, detail which does not lie near the surface closest to the objective appears clouded, like an object viewed through a plate of ground glass placed over it. The layer of matter between the interesting detail and the objective has an effect on the electrons similar to that produced by the irregularities of the ground glass on light. The resulting unsharpness is proportional to the 3/2 power of the thickness of the layer of matter intervening between the structure observed and the objective, as shown in section 19·6.

Both the chromatic aberration due to velocity losses within the object and the diffuseness resulting from the volume scattering of the electrons decrease with increasing applied voltage. The first is inversely proportional to the square of the voltage; the second, inversely proportional to the voltage itself. Figure 19·26 represents the change in the diameter of the circle of confusion due to both causes, with a change in the object thickness and for different voltages. For the usual operating voltages (about 50 kilovolts) and object thicknesses capable of leading to a satisfactory image the effect of chromatic aberration exceeds that of the volume scattering of the electrons. Only for very high voltages and thick specimens does the effect of volume scattering predominate. It is seen that the possibility of studying very thick specimens with the electron microscope is contingent on the application of very high accel-

[28] By Eq. 19·35*a*,

$$\frac{\Delta V}{V} \simeq \frac{10^{11}\rho x}{V^2} = 40\rho x$$

for $V = 50{,}000$ volts. ρx (ρ = density, x = thickness) is here the *mass density* of the object measured in g/cm^2. Thus, for $\rho x = 10^{-5}$ g/cm^2, $\Delta V/V = 4\cdot10^{-4}$. Accordingly, the diameter of the chromatic circle of diffusion for an object 1000 A.U. in thickness with unity density becomes, for an effective object aperture of $3\cdot10^{-3}$ radian and a focal length of 0.4 cm,

$$\frac{\Delta V}{V} \cdot \alpha f = 50 \text{ A.U.}$$

[29] See section 19·4.

erating voltages. An estimate of the resolutions that may be attainable under such conditions — taking account of the limitations of available electron-microscope objectives — is represented in Fig. 19·28.

4·8. Object Thickness and Contrast. The fact that the recognition of detail in objects may be limited, under certain circumstances, by inadequate contrast was first emphasized by L. Marton.[30] The eye cannot distinguish with certainty between two regions differing in

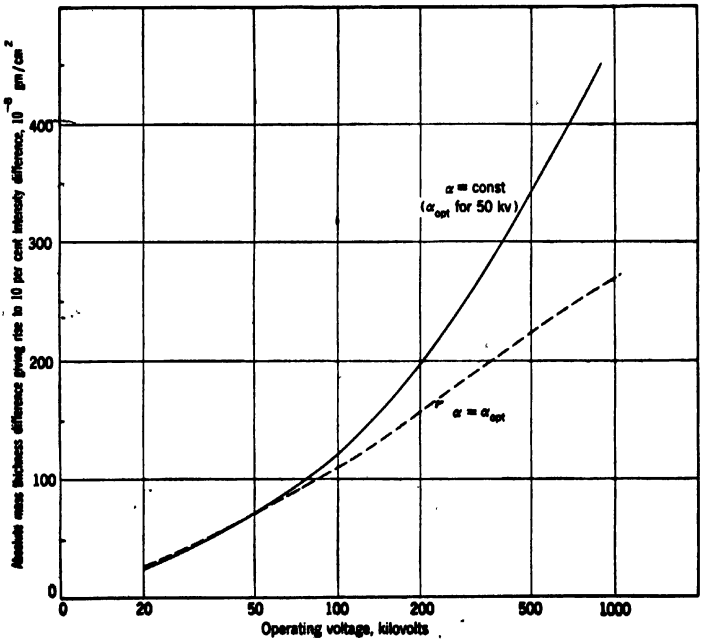


FIG. 4·26. Mass Thickness Differences for 10 per Cent Differences in Intensity as Function of the Accelerating Voltage.

brightness by less than a certain percentage. The latter, which varies from person to person as well as with the nature of the object, may arbitrarily be set at 10 per cent. Figure 4·26 shows a plot of the least difference in mass thickness that can be distinguished as a function of the accelerating voltage. One curve applies to the case in which the objective aperture is adjusted for optimum resolution at each voltage while the lens field is kept constant (f = 0.4 centimeter at 50 kilovolts). This least distinguishable thickness difference depends on the overall

[30] See reference 7.

thickness of the object only for very thick specimens.[31] If the physical aperture of the objective is large, the contrast ceases to be a simple function of the mass thickness, but depends also on the sizes, contours, and backgrounds of the objects, as is discussed in section 19·4. In general, the contrast will be found to be reduced by the employment of large apertures, particularly with relatively thick objects.

An interesting question arising in this connection concerns the possibility of detecting the smallest fragments of matter, that is, individual atoms.[32] Since the size of the image does not change as the object becomes substantially smaller than the limit of resolution of the microscope, the object size, as well as its thickness and density, are reflected in the relative intensity of the image and its background. Thus definite restrictions are placed by the limits of contrast recognition on the smallness of the particle — or, in the case of an atom, on the smallness of the atomic number — which can still be detected. Under actual circumstances an estimate indicates that a particle, to yield the desired contrast, would have to consist of, for example, at least 40 gold atoms, corresponding to a molecular weight of 8000.[33] Limitations of a similar character exist even when the particle is freely suspended in space and imaged by an aberration-free objective. Considering the interference between the electron wave scattered by the atom with the unperturbed wave in the image plane of the objective, Schiff shows that atoms must be heavier than the carbon atom to be detected. Only dark-field observation might permit the viewing or recording of the very light atoms, just as the light ultramicroscope reveals the presence of particles far below the limit of resolution of the light microscope. However, the practical difficulties occasioned by the scattering of the electrons in the support which is essential to keep the atoms in a fixed position, augmented by the unfavorable intensity conditions in dark-field observation, make this road appear unpromising.

4·9. Electrical and Magnetic Disturbances. The electrical and magnetic disturbances which may affect the quality of the final image are of two types. On the one hand, fluctuations in the voltage and current supplies alter the optical constants of the microscope; on the other, local electric and magnetic fields, unrelated to the lens and accelerating systems, may affect the electron paths directly.

The first, as has already been brought out in the discussion of the objective and projector, cause several kinds of chromatic aberration. In the objective these result in unsharpness throughout the image; in

[31] See section 19·2.

[32] See Hillier, reference 8, and Schiff, reference 9; see also section 19·8.

[33] See section 19·8.

the projector, in an unsharpness increasing linearly from the center outward. An increase in the coil currents is identical in effect with a reduction in the accelerating voltage. Von Ardenne[34] has utilized this effect for constructing a manual stabilizer (for use with the scanning microscope) which simply allows maintenance of the proper relative strength of the voltage and of the coil currents. It is more convenient and safer,[35] however, to stabilize the voltage and current sources individually as described in Chapter 7, the permissible fluctuations in the accelerating voltage and in the objective coil current being of the order of 0.01 per cent.[36]

The local disturbing fields which may affect the paths of the electrons originate either outside of or within the microscope. The effects of semiconducting or insulating contaminations within the instrument have already been mentioned in section 4·1. Care and cleanliness in assembling and operating the microscope as well as periodic cleaning of the critical parts, such as beam-limiting apertures, are the only remedy for these phenomena. The influence of external fields on the rays must be minimized, however, by shielding. The metallic construction of the microscope protects the interior adequately from external electrostatic fields. For magnetic fields, however, special provision must be made. Normally the entire ray path — at the very least the section between the objective and the projector lens — is surrounded by hollow cylinders of a material of high initial permeability, such as mu-metal or Permalloy (Fig. 14·11). This reduces magnetic-field fluctuations within the microscope to values which are too small to impair the resolution of the instrument.

As shown in section 16·10, only the component of the magnetic field normal to the optic axis need be considered, the disturbing effects of the axial component being small by comparison. The normal component of the magnetic field produces a deflection of the final image, which, for uniform field strength along the length of the beam, is proportional to this magnetic-field component and to the square of the distance between the object and the intermediate image. Furthermore, it is inversely proportional to the applied voltage. The relative sensitivity of the different parts of the ray path to the magnetic field is shown in Fig. 4·27. For equal distances between the objective and the intermediate image and between the projector and the final image, the effect on the first part of the ray path is more than that on the second part by a factor equal to

[34] See reference 10.

[35] In view of the different saturation properties of different lenses, resulting in variations in the relation $I = f(V)$ for which the focal length of the lens remains constant.

[36] See section 6·4.

the magnification of the projector. Again, the effect on the image of leaving the distance between the object and objective unshielded is smaller than that of leaving the distance between the objective and intermediate image unshielded by a factor equal to the reciprocal of the objective magnification — in reality, the coil casing and pole pieces act as an effective shield in the former region.

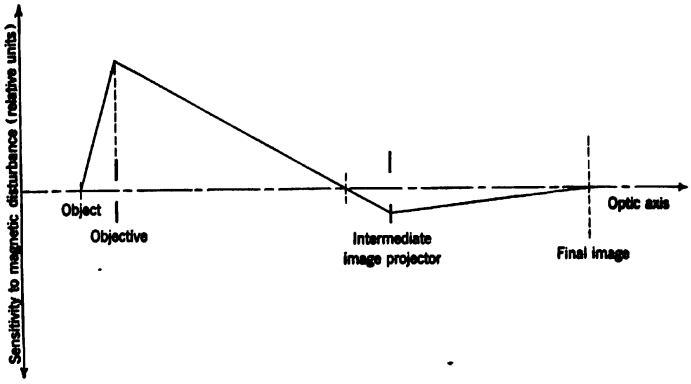


FIG. 4·27. Variation of the Sensitivity to Disturbance by Magnetic Fields along the Optic Axis of the Electron Microscope.

One source of magnetic-field disturbances is the fluctuation of the earth's magnetic field as well as of fields induced by the earth's field in ferromagnetic materials near the microscope. The fluctuation of the earth's field in periods of the order of a minute will rarely exceed 10^{-4} gauss. In the course of a day, it may attain a value as high as $7{\cdot}10^{-4}$ gauss. In magnetic storms, a fluctuation of an amplitude of $8{\cdot}10^{-3}$ gauss has been observed.[37] A second source of magnetic disturbance, which is of great importance, is the 60-cycle leakage fields of near-by electrical machinery. Von Ardenne has measured a field strength of $6{\cdot}10^{-4}$ gauss at a distance of 27 feet from a pair of 500-watt magnetic stabilizers. The effect of such leakage fields may be reduced by proper orientation of the machines and by increasing the distance between them and the microscope. In general, shields designed for static fields will be found to be even more effective for alternating fields because of the generation of eddy currents within them. At radio frequencies the microscope body itself constitutes an effective shield for magnetic as well as for electric disturbances.

[37] See von Ardenne, reference 11, pp. 108–111.

REFERENCES

1. F. Malsch, "Electron current density in the cold-cathode-ray discharge tube," *Arch. Elektrotech.*, Vol. 27, pp. 642–656, September 1933.
2. E. Schwartz, H. Strübig, and W. Paehr, "Beam formation in television tubes for projection," *Z. der Fernseh-A.G.*, Vol. 1, No. 1, pp. 5–13, August 1938.
3. E. Ruska, "A magnetic objective for the electron microscope," *Z. Physik*, Vol. 89, pp. 90–128, 1934.
4. J. Dosse, "Optical data of strong electron lenses," *Z. Physik*, Vol. 117, pp. 722–753, 1941.
5. M. v. Ardenne, "The testing of short-focal-length electron lenses — a simple method and its results," *Z. Physik*, Vol. 117, pp. 602–611, 1941; Vol. 118, pp. 384–388, 1941.
6. M. v. Ardenne, "On a universal electron microscope for bright-field, dark-field, and stereo operation," *Z. Physik*, Vol. 115, pp. 339–368, March 1940.
7. L. Marton, "Resolving power in electronic microscopy," *Physica*, Vol. 3, pp. 959–967, November 1936.
8. J. Hillier, "A discussion of the fundamental limit of performance of an electron microscope," *Phys. Rev.*, Vol. 60, pp. 743–745, 1941.
9. L. I. Schiff, "Ultimate resolving power of the electron microscope," *Phys. Rev.*, Vol. 61, p. 721, 1942.
10. M. v. Ardenne, "The electron scanning microscope: practical construction," *Z. tech. Physik*, Vol. 19, pp. 407–416, 1938.
11. M. v. Ardenne, *Übermikroskopie*, Julius Springer, Berlin, 1940.

CHAPTER 5

THE MAGNETIC ELECTRON MICROSCOPE

5·1. General Construction. In view of the fact that active research on the construction and improvement of magnetic electron microscopes has gone on continuously for well over a decade, it is not surprising that the instruments developed during this period differ in many details of construction even though they have in common all the basic elements described in the preceding chapter. Not all the designs have been equally successful. In particular, the earlier instruments tended to be uncertain in their results and difficult in adjustment. They demanded much skill on the part of the operator, so that their utility was confined to the laboratories of their origin. Only quite recently has it been possible to combine high performance with the compactness and simplicity of operation necessary in a successful commercial electron microscope. The extent to which this goal has been met in various modern microscopes together with their design features is the subject of this chapter. At the time of writing[1] two microscopes have been on the market long enough to enable their respective merits to be judged. They are the RCA Electron Microscope and the Siemens Supermicroscope. Some laboratory instruments and some commercial instruments which have been introduced recently have special features which will also be considered.

The Siemens electron microscope[2] is shown in Fig. 5·1. The microscope proper is a short column about three feet high mounted on a small table in front of a large stand containing pumping equipment and a lifting device. The cathode of the instrument is enclosed in a large protective casing at the top of the column. The necessary controls and meters are mounted on the front of the supporting table. The high-voltage generating equipment comprising a stabilizer, synchronous converter, high-voltage transformer, rectifier, and filter chain is set up in a separate room. This is essential here because of requirements of space and safety as well as to minimize the strength, at the instrument, of the low-frequency stray magnetic fields issuing from the electrical equipment. The lens coils are small but require water cooling to compensate for the high power

[1] Written in 1944.

[2] See v. Borries and Ruska, references 1 and 2.

dissipation. The coils, as well as the cathode filament, utilize storage batteries as their source of power. The filament battery is mounted on insulators within the high-voltage shield above the microscope. Control of the filament current and the grid voltage from the operating position is made possible by the use of insulated knobs projecting through this shield.

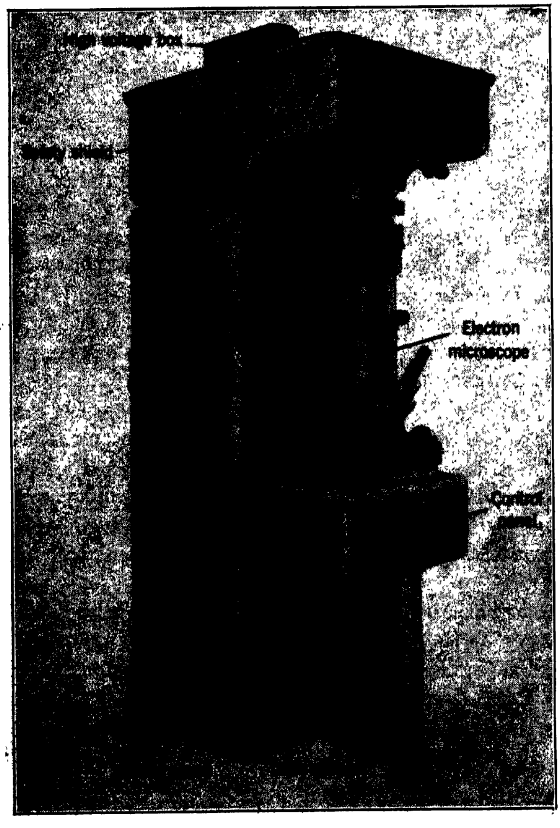


FIG. 5-1. The Siemens Electron Microscope. (v. Borries and Ruska, *Ergebn. d. exakt. Naturwiss.*, Vol. 19, pp. 237–322, 1940. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

Both the final image and the intermediate image are observed through any one of three large ports (for binocular viewing) in the final viewing chamber; pairs of prisms bring the intermediate image down to the

same viewing level as the final image. For the objective and projector pole pieces normally employed the magnification of the intermediate image ranges from 80 to 160, that of the final image from 4000 to 40,000. A focusing microscope with fourfold magnification is provided at one of the viewing ports. The projection-lens pole pieces can be exchanged, making low-magnification survey pictures possible, and the apertures in the instrument removed by using the lifting device incorporated in the rear stand. After letting air into the instrument, the handwheel shown at the left of the stand is actuated to lift the upper half of the microscope and to swing it sidewise, making the pole pieces of both the objective and the projector accessible. The illuminating system can be adjusted as a unit with respect to the rest of the microscope, whose elements remain in fixed relative position. The portion of the column between the objective and the projector is shielded with mu-metal; the iron construction of the microscope is deemed adequate magnetic screening for the remainder. Conveniently operated airlocks are provided for the insertion of both the object and the photographic plate.

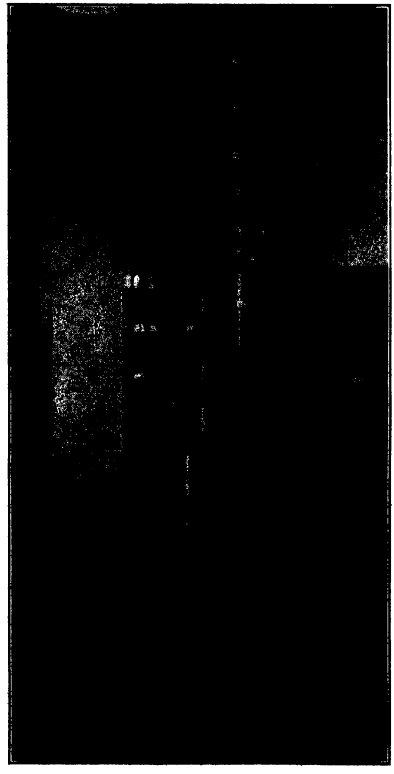


FIG. 5·2. The RCA Electron Microscope (Type B).

Ground grease joints are employed for transmitting motion to the inside of the vacuum chamber as well as at demountable junctions. Other seals are made with the aid of rubber gaskets. The evacuation is taken care of by a gas-heated mercury diffusion pump, connected to the microscope through a freezing-out trap and backed by a rotary oil pump.

The RCA electron microscope[3] also has been on the market for sev-

[3] See Zworykin, Hillier, and Vance, reference 3 and Hillier and Vance, reference 4.

eral years. A view of the Type B model, in its entirety, is shown in Fig. 5·2. Figure 5·3 is a sectional diagram of the same microscope. An important feature of this instrument is the inclusion of all the electrical equipment in the one microscope unit. The high-vacuum oil diffusion pump is also included in the main unit, whereas the forepump and an auxiliary pump are placed in a separate soundproof housing to eliminate the difficulties arising from mechanical vibration. This compact arrangement is made possible, for the most part, by the employment of radio frequencies in the generation of the high voltage, which greatly reduces the bulk of the transforming and rectifying equipment and, simultaneously, simplifies the problem of stabilization. The use of radio frequencies also greatly aids in eliminating the disturbing effects of stray magnetic fields from the high-voltage source.

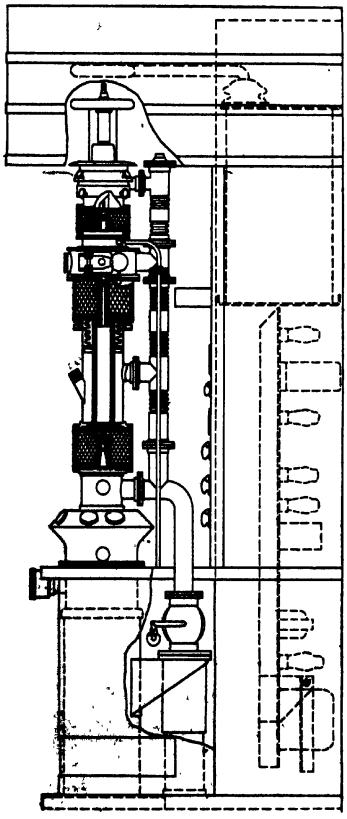


FIG. 5·3. The RCA Electron Microscope — Type B — (Vertical Section).

In the design of the Type B electron microscope an attempt has been made to simplify the operation and maintenance by avoiding the use of elements requiring periodic attention. Thus the need for storage batteries is eliminated by deriving all the power required for the instrument from the regular alternating-current line voltage. Both the coil currents and the filament current — in addition to the high voltage — are obtained from this source, electronic stabilization being provided to maintain the required constancy. The employment of an electrically heated oil diffusion pump renders a freezing-out trap unnecessary. In order to eliminate the need of servicing and the difficulties from internal surface contamination, ground grease seals are completely avoided. Instead, vacuum seals at demountable joints are maintained by rubber

and neoprene gaskets compressed between metal flanges. The displacement of parts, such as the object, the plate holder, the filament, the shutter, etc., within the vacuum is accomplished by the use of flexible metal bellows.

The dismounting and reassembly of the instrument are facilitated by its construction in a number of relatively short sections, each of which is seated on the preceding one, a neoprene gasket assuring a vacuum-tight junction. Where necessary, a special type of construction is used at these junctions, in which a metal-to-metal contact maintains rigidity and alignment and, in addition, minimizes the area of the sealing gasket which is exposed to the vacuum chamber (Fig. 5·4).

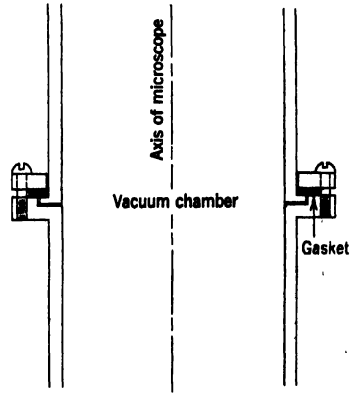


FIG. 5·4. Vacuum Seal Employed between Sections of the RCA Electron Microscope.

All control knobs, switches, and meters are placed on a panel in front of the operator, so that the necessary manipulations can be carried out conveniently from the seated observing position. Six round ports in the viewing chamber permit binocular observation of the final image by a group of operators. Further observing ports enable the operator to use the intermediate image for centering the illuminating beam and to observe the introduction of the specimen holder into the object stage. Easily manipulated airlocks speed the exchange of objects and photographic plates; provision is made for taking a number of pictures on a single 2- by 10-inch plate. A mu-metal shield consisting of three coaxial tubes protects the ray paths between the objective and the projector from disturbing magnetic fields.

The magnification obtained with the Type B electron microscope can be varied from 800 to 25,000. This variation is accomplished in two ranges, 800–3000 and 2000–25,000, corresponding to two possible positions of the object. Since the photographic plates normally employed in electron microscopy allow a subsequent optical enlargement by a factor of 10, the upper limit is sufficient to reveal the finest detail that can be resolved with the electron microscope, even if its resolving power should be made to exceed 10 A.U.

The Universal Electron Microscope[4] of von Ardenne is not a commer-

[4] See von Ardenne, reference 5.

cial instrument. However, it contains features that should be noted. It is shown, without the voltage supply, in Fig. 5·5. Source, condenser, objective, and projector are mounted in a column which, from the top of the high-voltage shield to the bottom of the camera, measures slightly over seven feet. The principal original feature of the instrument rests in the possibility of removing laterally all the vital elements, one by one, without dismantling the instrument as a whole: the condenser pole-piece unit with its exchangeable aperture, the objective pole-piece unit including the object stage, and the projector pole-piece unit. This facilitates exchanging the pole pieces and apertures, permitting even the replacement of the magnetic lenses by electrostatic lenses. Furthermore, the apertures can be centered and even exchanged in vacuum. To make all this possible, the pole pieces are offset vertically from the lens coils, which remain in the instrument when any of the several units are removed. In addition to the usual ports for observing the intermediate and final images on ordinary fluorescent screens, light microscopes are provided for observing the intermediate and final images on polished single crystals of activated zinc sulfide. The system used for viewing the final image is located just below the projection lens so that the actual electronic magnification on the screen is quite low and the intensity correspondingly high. The use of a large-aperture light microscope to view this image leaves its apparent brightness unaltered even at considerable optical magnification.

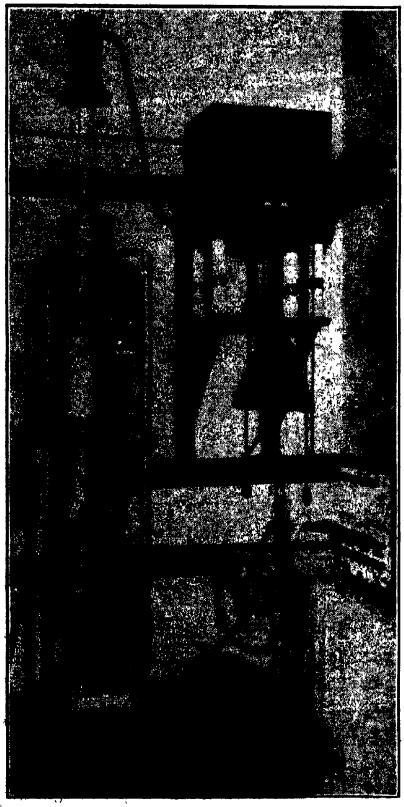


FIG. 5·5. The Universal Electron Microscope of von Ardenne. (v. Ardenne, reference 5.)

Permalloy shielding surrounds practically the entire ray path from the anode aperture to the viewing chamber. A special, large-diameter, shield is provided for the region about the objective. It rolls along four vertical columns and can be brought into place after necessary adjustments have been made on the objective unit. Finally, adjustment

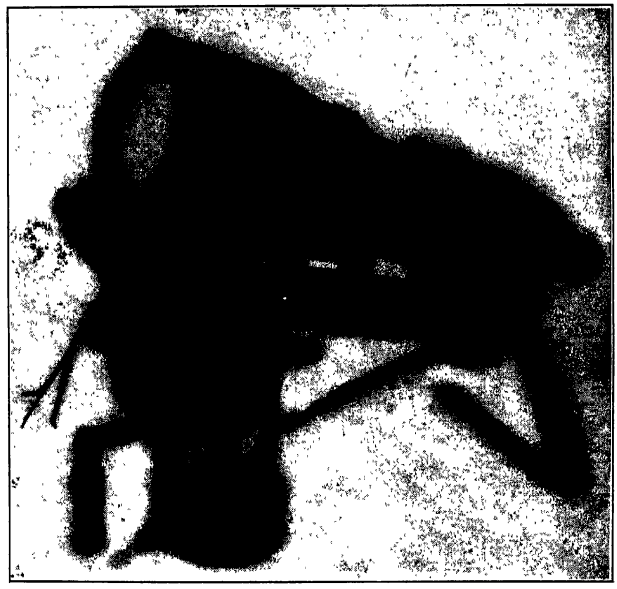


FIG. 5·6. Small Electron Microscope (Experimental Model). (Courtesy *J. Applied Phys.*, reference 6.)

screws on every unit make possible — in principle — perfect optical alignment. In view of the considerable height of the instrument and the very short focal length of the lenses (both objective and projector about one millimeter), direct magnifications up to about 500,000 can be achieved with this instrument — magnifications considerably in excess of the useful magnification even for a resolving power of 10 A.U.!

It is not surprising that, in an instrument of such extreme adjustability, some degree of convenience of manipulation has been sacrificed. For instance, in place of the more usual object airlock, two greased stopcocks, placed one above and one below the objective unit, must be closed and the whole objective pole-piece unit removed in order to exchange the object. As in the Siemens instrument, greased vacuum joints, sources of frequent trouble, are employed at several other points as well. The

good results obtained with the Universal Electron Microscope must in part be ascribed to the great skill of the operator.

The opposite tendency, namely, that of utmost simplification, both in construction and in operation, is represented by a compact electron microscope recently developed by the RCA Laboratories[5] (Figs. 5·6 and 5·7). Instruments of this type may reasonably be expected to play

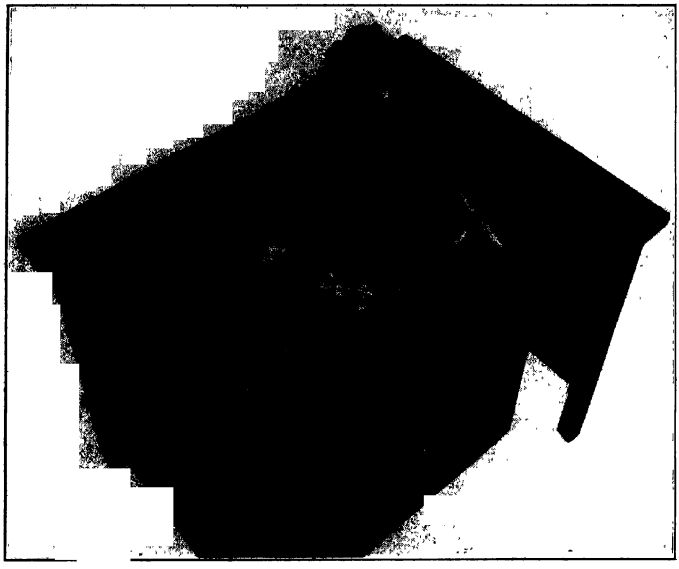


FIG. 5·7. Small Electron Microscope (Rear View, with Covers Removed). (Courtesy *J. Applied Phys.*, reference 6.)

a major role in broadening the distribution and hence the utility of the electron microscope. In the model shown the simplification is attained without sacrificing resolution. However, both the operating voltage and the magnification are kept at fixed values; the former is maintained at 30 kilovolts, the latter at either 500 (for survey pictures) or 5000. Since experience has shown that the majority of electron-microscopic investigations are carried out with fixed operating voltage and magnification, this does not, normally, represent a serious drawback. Furthermore, if the usual fine-grained photographic materials (with a resolution, for example, of 100 lines per millimeter) are utilized, an electron-optical magnification of 5000 together with a light-optical enlargement by a factor of 16 will reveal image detail with a resolution of 20 A.U. Fixing

[5] See Zworykin and Hillier, reference 6.

the magnification at a low value leads to the advantages of a short microscope column and of bright and large image fields.

Unlike the larger instruments discussed here, this microscope is not arranged vertically. The shortness and rigidity of the microscope column permit its orientation to be selected simply on the basis of the convenience of the observer. It is thus placed, with its axis inclined at an angle of about 20 degrees to the horizontal, on a small-sized stenographic desk, so that the transmission-viewed screen is approximately normal to the line of sight of the seated observer.

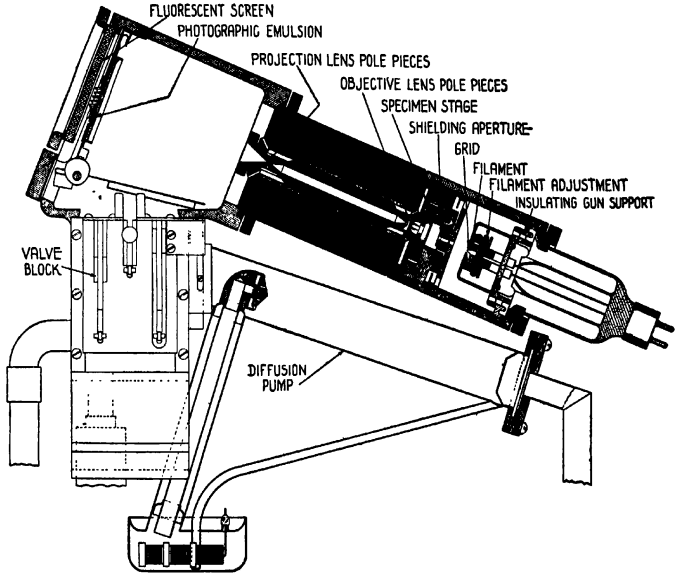


Fig. 5·8. Small Electron Microscope (Vertical Section, Experimental Model). (Courtesy *J. Applied Phys.*, reference 6.)

The interior arrangement of the microscope is evident from the sectional drawing in Fig. 5·8. It will be noted that, in place of a condenser lens, a focusing aperture with a variable bias potential serves to control the convergence and intensity of the illumination on the specimen; a shielding aperture screens the outer portions of the specimen support from bombardment and thus prevents overheating. The objective and projection lenses are formed by a single magnetic circuit, thus requiring only a single lens coil and power supply. The final image is formed on a highly efficient fluorescent transmission screen or a photographic plate,

2 by 2 inches in size, which may be swung in front of it. It is clear that the image may also be photographed externally, at the expense of a greatly lengthened time of exposure and possible loss of detail.

The vacuum within the instrument is created by a high-speed, single-stage oil diffusion pump backed by a rotary oil pump. No airlocks are provided since the time required for evacuating the complete microscope is only 1½ to 3 minutes, depending on the length of time it has been left open.

The power supplies of the microscope, visible at the rear of the desk in Fig. 5·7, are constructed on the same general principles as those of the Type B electron microscope. Their complexity and space consumption are, however, greatly reduced by the lower operating voltage, the smaller number of circuits, and the diminished demands on variability.

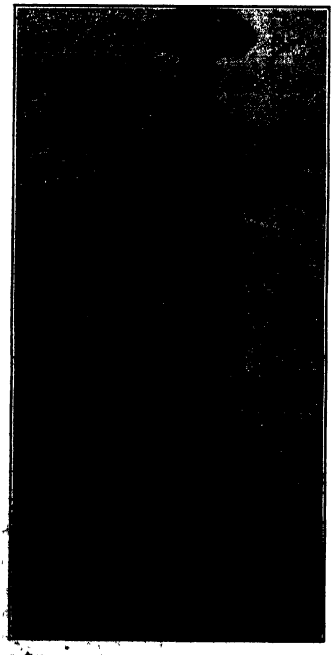


FIG. 5·9. Early Magnetic Electron Microscope. (Ruska, reference 7.)

Before proceeding with a more detailed description of the several parts of the electron microscopes discussed above, it is worthwhile to refer briefly to a number of other, largely earlier, magnetic electron microscopes which differ from those described in certain interesting details.

The first high-magnification two-stage electron microscope for the study of specimens by transmitted electrons was constructed by Ruska[6] in 1933. It was derived largely from the cold-cathode-ray oscillograph and is similar, except for two major differences, to its successor, the latest Siemens microscope. The employment of a cold-cathode gas-discharge tube as electron source and the absence of airlocks for exchanging the object and the plate constitute these differences (Fig. 5·9). A series of objects could be placed in apertures drilled in a horizontal, eccentrically mounted cogwheel. The cogwheel was rotated by a pinion, manipulated with the aid of a greased joint, which brought the objects in succession in front of the objective. The final image, formed on a fluorescent screen, was photographed through a window.

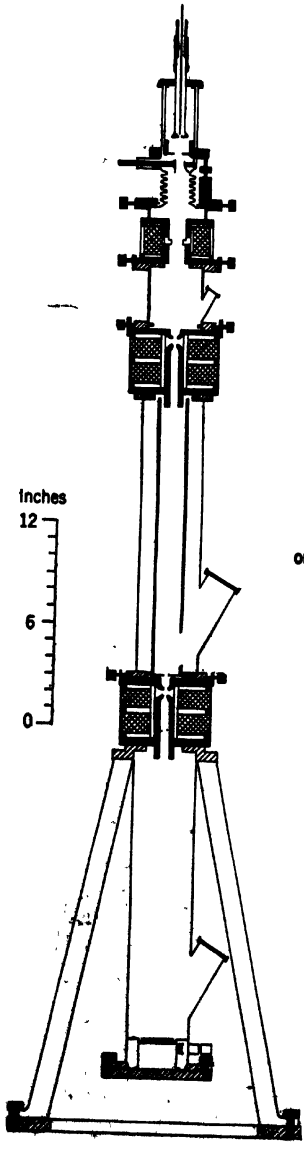
[6] See reference 7.

In discussing the performance of this instrument Ruska pointed out the advisability of replacing the gas discharge tube by a steadier, hot-cathode, electron source.[6] He also recommended the use of an object airlock and the direct recording of the electron image on a plate inserted into the vacuum. These latter improvements were incorporated in an electron microscope constructed at the University of Brussels by Marton,[7] who had the distinction of being the first to apply an instrument of this type to the examination of biological specimens. At a later date, an electron microscope differing in various details of construction from its predecessors was built in England by the Metropolitan-Vickers Electrical Company for Martin, Whelpton, and Parnum.[8]

Meanwhile, Driest and H. O. Müller,[9] working with Ruska's instrument modified for internal photography, had demonstrated that the electron microscope was, in fact, capable of a resolving power exceeding that of the best light microscope. Results obtained with more recent instruments removed all doubt regarding this matter. A highly successful microscope design, attractive in its simplicity, is represented in the diagrammatic cross section of the electron microscope constructed by Prebus and Hillier[10] at Toronto (Fig. 5·10). Like the most modern designs, it possesses a hot cathode as source and magnetic shielding between the objective and the projector. Although no airlocks were provided originally, a plate airlock was added later.[11] The vacuum junctions are largely greased joints. Provision is made for translating the several parts of the microscope relative to each other so as to make good alignment possible. A similar instrument, which has also yielded some fine micrographs, has been constructed by the Eastman Kodak Company.[12]

The microscope shown in section in Fig. 5·11 is an earlier construction developed at the RCA Laboratories.[13] It deviates greatly in its mechanical construction from all other instruments, with the possible exception of Martin, Whelpton, and Parnum's. Great stress is placed on rigidity. The whole optical system, including the coils sealed in copper cans, is mounted inside heavy, large-diameter brass cylinders. These are broken only by the object airlock and the viewing ports for the final image. A periscopic arrangement is provided for observing the intermediate image through the same ports. An interesting feature of this microscope consists in the introduction of the object, placed between two flat, apertured

[7] See reference 8.
[8] See reference 9.
[9] See reference 10.
[10] See reference 11.
[11] See Burton, Hillier, and Prebus, reference 12.
[12] See Hall and Schoen, reference 13.
[13] See Marton, Banca, and Bender, reference 14, and Marton, reference 15.



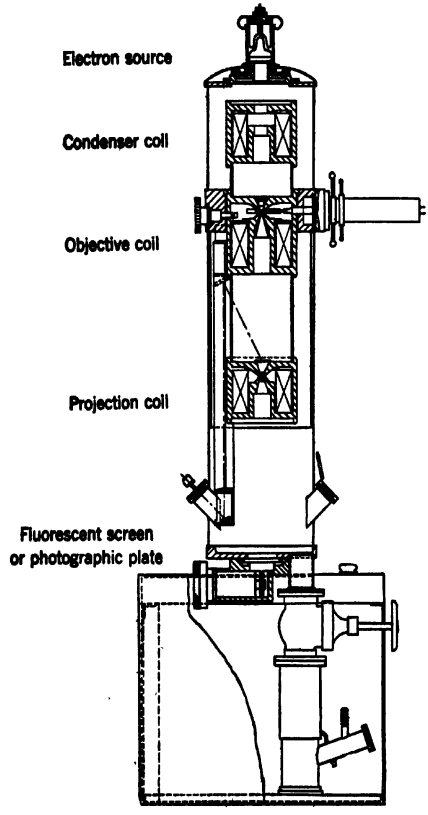


FIG. 5·11. RCA Electron Microscope, Mode A. (Courtesy *Phys. Rev.*, reference 15.)

FIG. 5·10. Electron Microscope of Prebus and Hillier (Toronto). (Courtesy *Can. J. Research*, reference 11.)

brass blades, into the region between the two pole pieces of the objective. The object film is thus mounted on a piece of wire mesh, which, within wide limits, may be arbitrary in size. Though it is capable of yielding pictures of high resolution, the great difficulty experienced in aligning and adjusting this instrument, as well as the relatively large amount of space which it occupies, render it unsuitable for general research purposes.

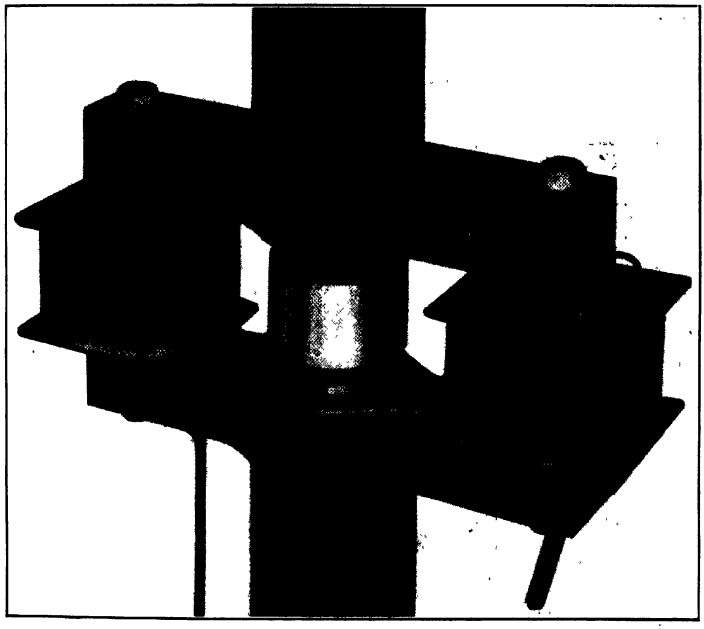


FIG. 5·12*a*. Magnetic Yoke Lens: External Appearance (Kinder and Pendzich, reference 16. By permission of the Alien Property Custodian in the public interest under License No. A-563.)

A magnetic electron microscope which differs from those described so far primarily in the method of energizing its lenses is the yoke-lens microscope of Kinder and Pendzich.[14] In place of the usual axially symmetric arrangement of the lens coils, the latter are wound on the arms of a double yoke on either side of the pole pieces. The external appearance and a diagrammatic cross section of one of the lenses are shown in Fig. 5·12, *a* and *b*, respectively. The advantage of this arrangement lies

[14] See reference 16.

in the case of both exchanging and cooling the coils. Difficulties arising from the asymmetry of the relatively large stray fields of this type of lens are overcome by attaching soft-iron cylinders on either side of the pole pieces, which shield the region near the lens (Fig. 5·12*a*). It is also possible to replace the lens coils with permanent magnets with a minimum of effort. The focal lengths achieved with yoke lenses are of the same order as those obtained with the more conventional axially symmetric lenses.

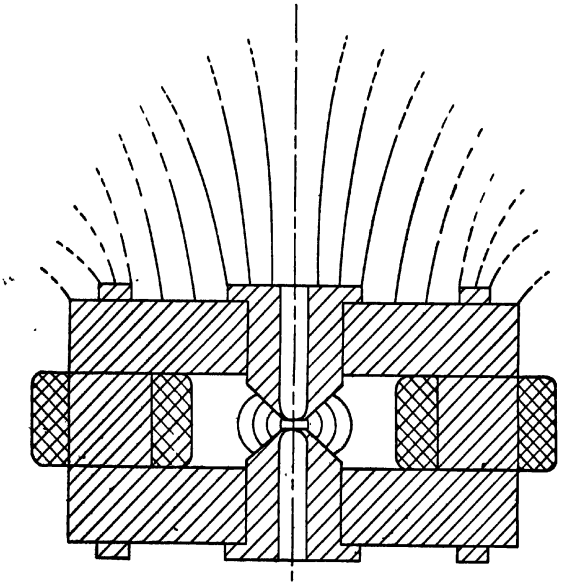


FIG. 5·12*b*. Magnetic Yoke Lens: Section, without Symmetrizing Cylinders; Leakage Field Represented on One Side of Lens Unit Only. (Kinder and Pendzich, reference 16. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

An electron microscope which employs permanent-magnet lenses has also been constructed by von Borries, Ruska, Krumm, and Müller.[15] Here the lens coils of an ordinary Siemens microscope and their outer casing were replaced by thirty-two parallel bar magnets. With this arrangement, slightly over one-third of the maximum refractive power attained with the coils was realized. Thus the focal lengths of the objective and the projector were 7.5 and 2.7 millimeters, respectively. Particles with centers separated by 160 A.U. were found to be resolved.

[15] See reference 17.

Focusing must here be accomplished by changing the beam voltage, by displacing the object relative to the objective, or by shunting the magnetic field.

Satisfactory results with permanent-magnet excitation of the lenses have also been obtained with the small electron microscope already described.[16] It is here an easy matter to replace the lens coil by a set of bar magnets. In this manner a total electron-optical magnification of 1500 is obtained. It should be noted that the simplification attained by rendering the power supply for the lens coil unnecessary is here compensated by the need of a new focusing control, requiring a variable high voltage, a precision adjustment of the position of the specimen along the optic axis, or a mechanical means of varying the strength of the magnetic field. For this reason permanent-magnet excitation has not, to any extent, replaced current excitation of magnetic lenses in electron microscopes.

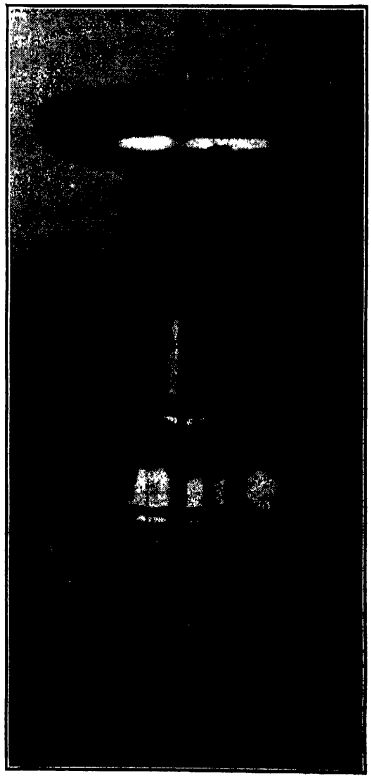


FIG. 5·13. Beam-Forming System of Type B Electron Microscope.

5·2. The Electron Source. The purpose and the general arrangement of microscope electron sources have been outlined in the preceding chapter. Their construction varies, however, in the several instruments in a number of important details.

A view of the complete beam-forming system of the Type B electron microscope is shown in Fig. 5·13. Above the corona ring at the top of the gun are visible the connections of the filament and, at the center, the knurled nut which permits the raising and lowering of the filament within the highly polished cathode cylinder. The latter has a ⅛-inch aperture at the bottom. Below it is the concave copper anode cup with

[16] See Zworykin and Hillier, reference 6.

a ¼-inch aperture, whose support is provided with a set of large pump holes for accelerating the evacuation of the cathode region. Next to the cathode corona shield the adjustment screws for the two alignment blocks of the electron gun are visible. Below them a port makes possible the observation of the intensity distribution of the beam on an inclined apertured fluorescent screen. The assembly is terminated by the condenser-coil casing.

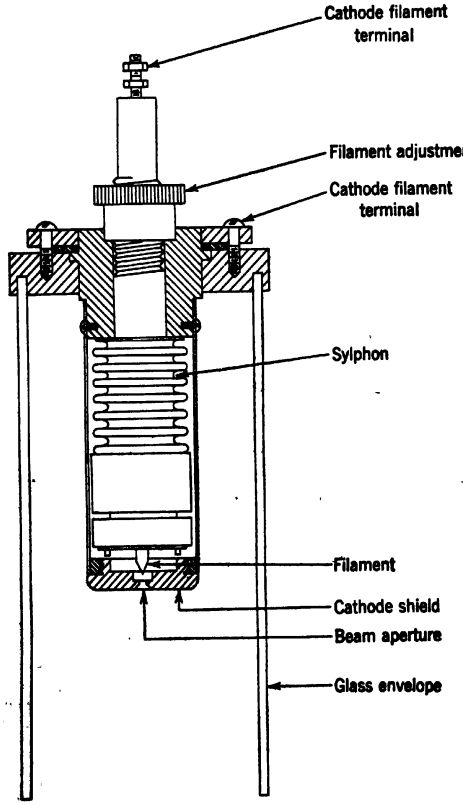


FIG. 5·14. Cathode Assembly of Type B Electron Microscope (Section).

The mounting of the filament in the cathode cylinder is visible in greater detail in Fig. 5·14. It is seen how a rotation of the knurled head permits the adjustment of the filament mount, expanding or contracting the metal bellows. The whole filament and cathode assembly can be

withdrawn after the vacuum has been broken by loosening the set of screws at the top, which compress a neoprene gasket between two metal flanges. The bottom flange forms part of a brass ring sealed with picein to the glass insulator. The triangular tungsten filaments themselves are welded to sturdy metal stems mounted on mica disks and are inserted into the centering mount after the cathode cylinder has been slipped off.

A second picein seal joins the glass insulating cylinder to the anode assembly and the beam direction adjustment block. A vacuum-tight connection between the latter and the condenser chamber is made by means of a second metal bellows.

The underside of the beam direction adjustment block has a circular rim which rests on a spherical surface. The center of curvature of the latter coincides approximately with the tip of the cathode filament. The rim may be displaced along this surface with the aid of the upper set of adjustment screws visible in Fig. 5·13. This motion serves to orient the beam relative to the optic axis of the condenser without varying the position of the source. The spherical surface itself, in turn, forms part of the beam position adjustment block, which rests on the plane surface of the condenser coil chamber. With the aid of the second set of screws it may be translated horizontally, displacing the beam as a whole in any direction normal to the optic axis. The two sets of adjustment screws thus permit a complete alignment, in both position and direction, of the beam relative to the condenser. A further set of screws is provided for the horizontal displacement of the condenser, together with the gun structure mounted above it, relative to the object and the objective.

The von Ardenne and Siemens instruments similarly employ easily exchanged hairpin cathodes, which, here, are precentered in porcelain plugs. The arrangement of the sources, however, differs from that of the Type B microscope in several details. Thus, in place of providing a possibility of adjusting the height of the filament in the cathode cylinder, the cathode cylinder is given a variable negative bias. The vacuum seals, as previously mentioned, are largely greased joints. Furthermore, the adjustments for aligning the beam deviate materially in the two instruments. In the von Ardenne microscope tilting and horizontal displacement are provided for both the gun and the condenser coil. The tilting, however, is accomplished simply by the motion of vertical screws acting against atmospheric pressure compressing rubber rings placed between successive sections of the microscope. It is thus accompanied by a simultaneous horizontal displacement of the source, increasing the difficulty of alignment. In the Siemens instrument, on the other hand, the cathode system and the condenser form a rigid unit, as shown in Fig. 5·15. A large measure of reliance is placed on the exact precentering of the

cathodes in their mounts and on the maintenance of their original position during operation. Both horizontal displacement and tilt are provided for the complete gun-condenser unit. The center of the tilt rotation is placed at the object instead of at the source. This arrangement

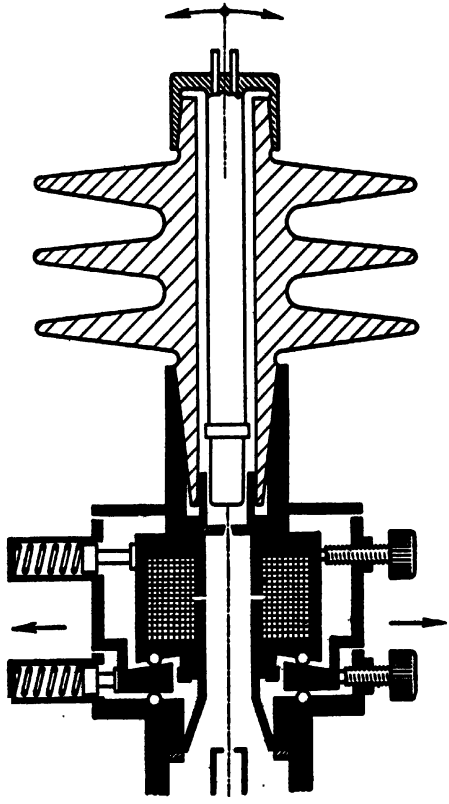


FIG. 5.15. Beam-Forming System of Siemens Electron Microscope (Section). (v. Borries and Ruska, *Ergebn. d. exakt. Naturwiss.*, Vol. 19, pp. 237–322, 1940. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

is advantageous primarily for passing from bright field to dark field, this requiring merely a displacement of the upper (tilt-adjustment) screws. As seen in the figure, a rubber collar, clamped at both ends between metal rings, maintains the vacuum-tight connection between the gun-condenser unit and the remainder of the microscope.

In the small electron microscope the cathode structure consists of a small removable capsule which includes the focusing aperture and an adjustable filament support for the hairpin filament (Fig. 5·16). When a new filament is inserted, it is first centered visually with respect



FIG. 5·16. Experimental Cathode Assembly of Small Electron Microscope. (Courtesy *J. Applied Phys.*, reference 6.)

to the focusing aperture, then placed in the microscope and heated in vacuum to the normal operating temperature, and finally removed and recentered. The resulting position of the filament is normally maintained throughout the rest of its life. A slight amount of tilting adjustment is provided to perfect the alignment of the illumination with the optic axis of the instrument. The intensity and convergence of illumination are controlled by a variable potential of 0 to -75 volts placed on the focusing aperture. This brings the virtual image of the source to a point close to the cathode assembly, leading to an angular aperture of illumination at the specimen between 10^{-3} and 10^{-4} radian. Since, in the absence of a condenser lens, the beam cross section at the specimen is relatively wide (about one millimeter) even under the condition of maximum beam concentration, a shielding aperture (Fig. 5·8) is provided which restricts the illumination to an area only slightly greater than that which can be observed with the microscope at any one time. Without this precaution the energy absorbed by the wires supporting the specimen may be great enough to cause the overheating and hence the destruction of the specimen.

5·3. The Condenser, Objective, and Projector Lenses. The magnetic electron microscope usually contains three magnetic lenses. The general arrangement of those in the Type B electron microscope is

evident from Fig. 5·3. In view of the weaker fields required, the condenser lens is not equipped with pole pieces and is provided with a smaller coil than the two other lenses. It is located approximately halfway between the source and the object, so that, when adjusted to give maximum intensity of illumination, it forms an image of the source with

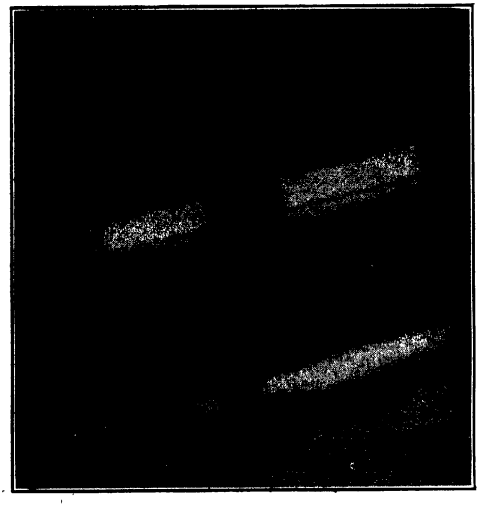


FIG. 5·17. Objective Pole Pieces and Aperture Insert (Type B Electron Microscope).

unity magnification on the object. The objective and projector coils are seen to be made relatively long and of small inner diameter, resulting in very low power loss. Provision is made for the insertion of limiting apertures in both the condenser and the objective. Centering screws for the lateral displacement of the objective and the structure above it are provided to make possible alignment with respect to the projector coil.

The pole pieces have the general character of those shown in Fig. 4·17. They form a single cylindrical plug, which fits exactly into the pole-piece support. Figure 5·17 shows a typical pair of objective pole pieces together with the aperture system, which screws into the bottom pole piece. The whole unit may be lifted out of the support and out of the microscope through the object chamber by means of special tongs gripping the circular groove near the top of the upper pole piece, the object airlock mechanism having been removed beforehand. The aperture system consists of a long, narrow brass tube suspended by means of a

brass diaphragm in a wider brass tube. At the top of the narrow brass tube is a cap, which holds the aperture in place. The aperture itself consists of a hole, normally about $\frac{1}{1000}$ inch in diameter, drilled into a small disk of sheet copper. Four centering screws at the bottom of the wide tube displace the aperture laterally by tilting the narrow tube relative to the supporting diaphragm. The centering of the aperture in the pole pieces is verified by rotating the pole pieces, whose outer surface is machined with as great precision as its inner surfaces, in a V-block, observing the aperture with a light microscope. The centering screws are then adjusted until the aperture appears to remain stationary during the rotation. For many electron-microscopic investigations — particularly such as are concerned with relatively thin specimens — the use of an objective aperture is superfluous. In such cases the omission of the aperture system is advantageous, since contamination of the objective aperture, with consequent charging effects, is one of the commonest sources of defective images.

Inasmuch as precise focusing of the objective is essential for the obtaining of good images, three control ranges are provided for the objective current, a coarse adjustment of 10 positions, a fine adjustment of 21 positions, and a continuous vernier adjustment. Normally the magnification produced by the objective alone is 100, corresponding to a focal length of 4 millimeters. With the specimen placed at the higher of the two possible positions, suitable for obtaining survey pictures, the objective magnification is about a third of this. The projector magnification can be varied from 25 to 250 (focal lengths from 16 to 1.6 millimeters) in 20 steps. A continuous control is provided for the condenser coil current. Any one of the three coil currents, furthermore, may be read on a meter installed for this purpose.

In general character, the lenses of the Siemens and von Ardenne instruments are similar to those of the Type B electron microscope. In the Siemens instrument both the focal lengths (down to 2 millimeters and 1.1 millimeters for the objective and the projector, respectively) and the image distances (32.5 and 26.5 centimeters) are shorter and yield slightly higher total magnification. If survey pictures of a magnification of the order of 1000 are desired, the upper part of the microscope must be swung aside and the pole pieces of the projector removed.

In von Ardenne's electron microscope short focal lengths are combined with long image distances, leading to very high magnifications, as mentioned in the initial section. The most interesting feature, however, is the mechanical arrangement of the pole-piece units and the adjustable apertures. Figure 5·18 shows the complete objective pole-piece unit removed from the microscope. At the left are seen the pole pieces proper

with one of the three pushers serving to displace the object within the upper pole piece. The major part of the remainder of the assembly is covered by the clamping ring, which compresses a rubber gasket between

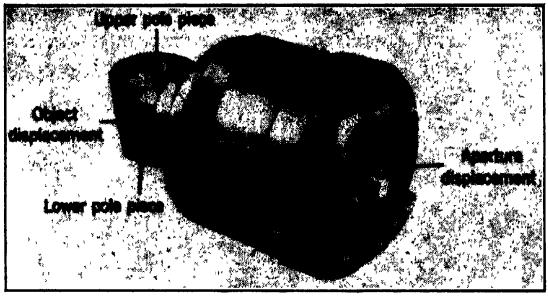


FIG. 5·18. Objective Unit of von Ardenne's Universal Electron Microscope. (v. Ardenne, reference 5.)

the pole-piece unit and a flange of the microscope body so as to secure a vacuum-tight fit. At the extreme right a micrometer head is visible, which is used to displace radially between the pole pieces a tongue bearing a succession of fine apertures. In addition, one of the two bolts for lateral centering of the aperture tongue may be seen at the left of the

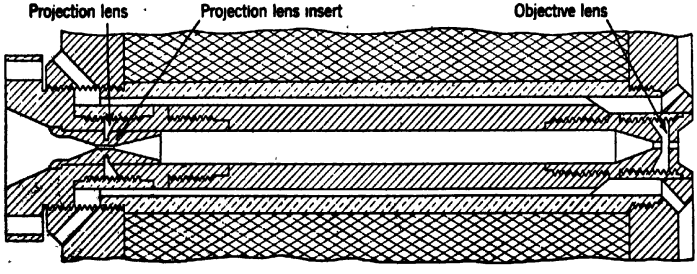


FIG. 5·19. Experimental System of Small Electron Microscope (Section).

micrometer head. A metal bellows permits the displacement of the aperture tongue without breaking the vacuum. Both the condenser and the projector units are similar to the objective unit, each having an adjustable aperture system.

A detailed diagram of the objective and projector lens system of the small electron microscope is shown in Fig. 5·19. With the fluorescent screen removed, the entire pole-piece system of the two lenses

may be slipped out of its support after a few counterclockwise turns of the head at the left. With comparable ease, the pole-piece insert of the projection lens may be removed. With this insert in place the system has a magnification of 5000; without it, one of approximately 500. Since all the parts of the two lens systems fit into a single accurately machined unit, adjustments for the alignment of the two lenses become quite superfluous.

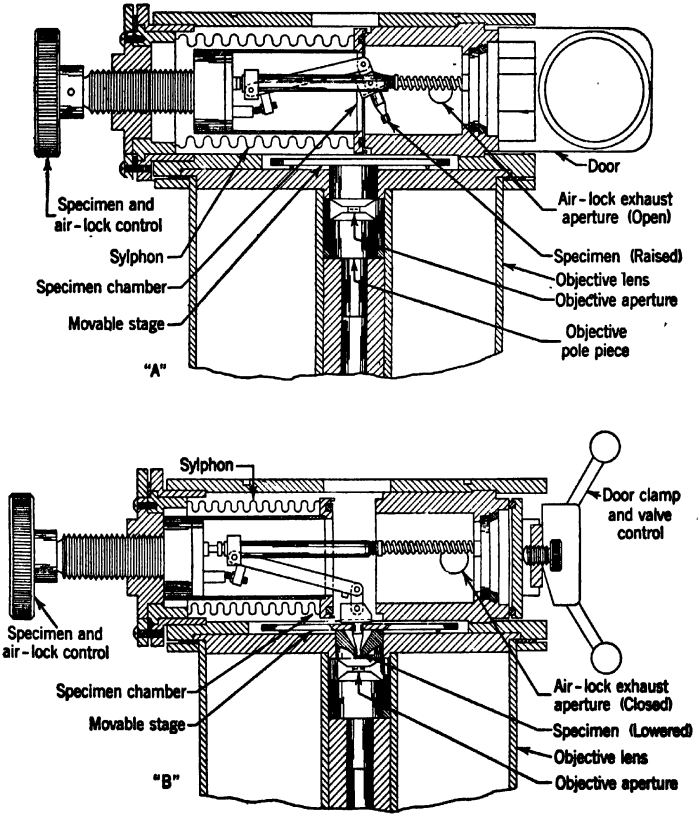


FIG. 5·20. Specimen Chamber Mechanism of the Type B Electron Microscope (Section): (*A*) Chamber Open; (*B*) Chamber Closed with Object in Place.

5·4. Object Chamber. The object chamber must perform two functions. On the one hand, it must provide for the insertion and withdrawal

of the specimen with a minimum loss of time between successive observations of different specimens; on the other hand, it must make possible the controlled lateral displacement of the object during observation. In instruments having large volumes the use of an airlock becomes necessary if the first requirement is to be satisfied.

Many ingenious solutions have been evolved for accomplishing both purposes with the greatest ease and certainty. The specimen chamber mechanism of the Type B electron microscope is shown in Fig. 5·20, *A* and *B*. In Fig. 5·20*A* the airlock is open and the slide of the object cartridge (Fig. 5·21) has been inserted with a pair of forceps through the open door into the cartridge support. The airlock control at the left has been screwed in until the flange connected to the metal bellows has made vacuum-tight contact with the rim of the brass cylinder forming the outer portion of the specimen chamber. Thus the specimen chamber is sealed off from the rest of the microscope. After the insertion of the specimen cartridge the door is closed. A rotation of the door clamp through a 180-degree angle shuts off the connection to the outer atmosphere and clamps the door. A further rotation, through 90 degrees, connects the chamber to the auxiliary pump, resulting in its speedy evacuation. After the door clamp has been turned back 90 degrees into its neutral position, where both the auxiliary pump and the air connections are sealed, the airlock control is turned counterclockwise, opening the inner gate and, by means of the indicated system of levers, dropping the specimen cartridge into the object stage (Fig. 5·20*B*). The microscope is now ready for observation. Depending on the position in the cartridge of the cartridge cap with the object support (Fig. 5·21), images of high or of low magnification are obtained.

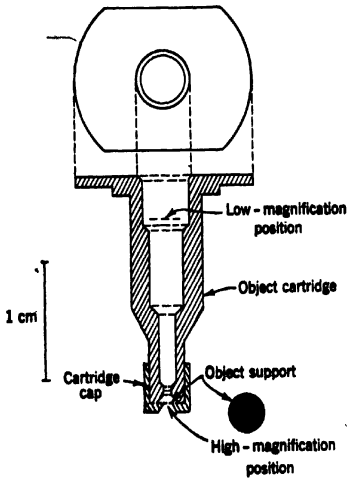


FIG. 5·21. Specimen Cartridge, Specimen Disk, and Cap (Type B Electron Microscope).

The stage itself rests on three ball bearings on the upper surface of the objective coil casing. It is displaced in two mutually perpendicular directions by a pair of screws acting against springs parallel to the screw

motion, as shown in Fig. 5·22. Figure 5·23 gives a full view of the object chamber assembly.

The airlock of the Siemens instrument, which differs greatly from that just described, is shown in Fig. 5·24, *a* and *b*. The cartridge, with the

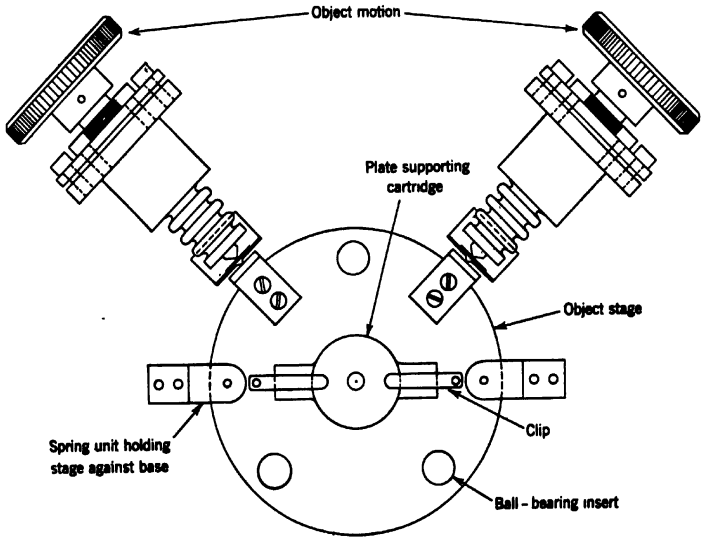


FIG. 5·22*a*. Object Stage of the Type B Electron Microscope: Top View.

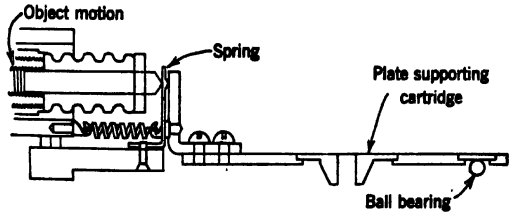


FIG. 5·22*b*. Object Stage of the Type B Electron Microscope: Vertical Section.

specimen mounted on a diaphragm at the narrow end, is inserted in the bore of a stopcock whose axis is at right angles to the axis of the instrument. Then the stopcock is rotated, so that the bore is aligned with the beam, the small volume of air within the bore diffusing into the remainder of the microscope. Finally, a pinion causes a plunger to

push the cartridge down toward the objective, until a flange on the object holder comes to rest on the upper pole piece, being pushed against it by a spring in the interior of the cartridge. The object motion is produced by two right-angle levers actuated by vertical columns and opposed by a pair of springs, the whole section of the microscope above the



FIG. 5-23. Specimen Chamber and Objective Unit of the Type B Electron Microscope.

object coil being displaced. This is practicable here, since the object area which can be explored is much smaller than in the Type B microscope. The object supports employed in both the Siemens and the von Ardenne microscopes are disks with apertures 0.05 to 0.1 millimeter in diameter, whereas in the Type B instrument the area of object supporting screen which can be examined is approximately one millimeter in diameter. The direct pressure contact introduced by von Ardenne between the object mount and the pole pieces is intended to minimize the effect of mechanical shock on the image.

As mentioned before, von Ardenne's electron microscope does not have a true object airlock. The whole objective unit is removed after the portion of the microscope column occupied by it has been sealed off

from the remainder of the microscope by two stopcocks. The object mount is pressed by a spring against a plane pole-piece surface (Fig. 5·25). For the displacement of the specimen, it is pushed in any one of three directions differing by 120 degrees, the pushing rod being retracted

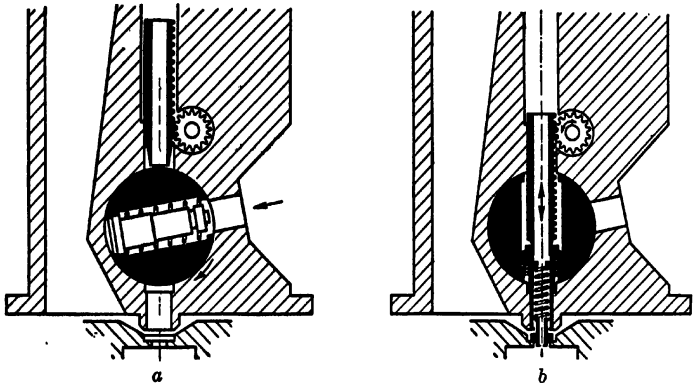


FIG. 5·24. Object Airlock of the Siemens Electron Microscope (Section): (*a*) Insertion of Specimen Cartridge; (*b*) Specimen in Place for Observation. (v. Borries and Ruska, *Ergebn. d. exakt. Naturwiss.*, Vol. 19, pp. 237–322, 1940. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

after every operation. By this method von Ardenne seeks to exclude any possibility of motion of the object relative to the objective during exposure; for example, due to the thermal expansion of the displacing mechanism. This precaution may be significant in the long exposures

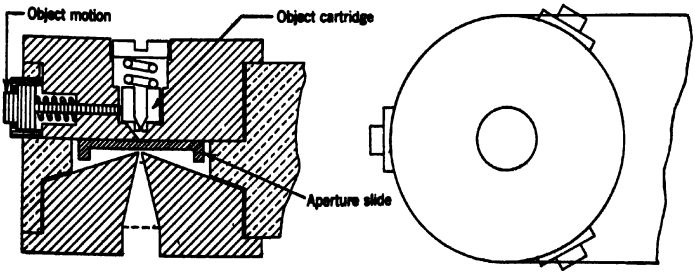


FIG. 5·25. Object Displacement Mechanism. (v. Ardenne, reference 5.)

required by dark-field pictures with the very high direct magnifications of which only the von Ardenne instrument is capable.

In the small electron microscope shown in Fig. 5·6 the complete

object stage, including the object mount, is removed from the side of the instrument whenever the specimen is to be exchanged (Fig. 5·26). The stage proper consists of a movable plate mounted on ball bearings to permit motion normal to the axis of the instrument. The motion is

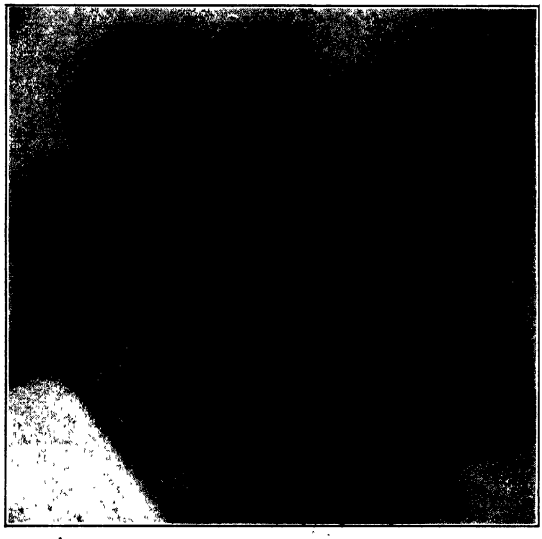


FIG. 5·26. Experimental Object Stage of the Small Electron Microscope. (Courtesy *J. Applied Phys.*, reference 6.)

controlled by two adjustment screws threaded into the body of the specimen stage, which work against a helical spring under tension placed between them. The vacuum is maintained by seals of the packing gland type. The specimen rests on a tubular projection on the plate and is secured by a cap. A specimen area approximately 2.5 millimeters in diameter may be explored with this microscope.

5·5. Plate Chamber. The plate chamber also has a number of functions to perform. First of all, it should permit the insertion of photographic plates into the microscope without unduly delaying operations. Second, provision must be made for removing the cover of the plate holder after it has been inserted into the microscope, so that the exchange of plates can take place in a lighted room. Finally, a shutter has to be provided for exposing the plate after it has been brought into place and after the interesting portion of the object has been selected and focused.

The plate-chamber mechanism of the Type B electron microscope is shown in Figs. 5·27*A* and *B* in section; Fig. 5·28 is a photograph of the complete assembly. In the first figure (5·27*A*) the plate carriage has been racked all the way to the left by means of the pinion actuated by

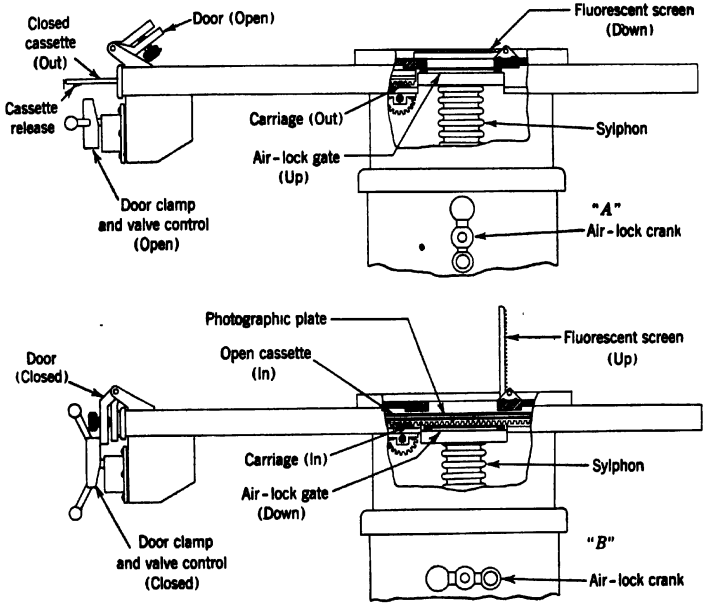


FIG. 5·27. Plate Chamber Mechanism of the Type B Electron Microscope (Section): (*A*) Plate Being Inserted; (*B*) Plate Being Exposed.

the graduated head in front of the table in Fig. 5·2. The closed plate holder has been partly withdrawn from the plate chamber through the open door. The inner gate, between the microscope interior and the plate chamber, has been sealed by a metal door provided with a rubber gasket. The door is pushed against the gate by a counterclockwise rotation of the airlock crank. A metal bellows seals the plate chamber against the outside air.

After inserting a holder with a fresh plate, the door is closed and the door clamp rotated through 270 degrees, causing the chamber to be evacuated by the auxiliary pump. In about a minute, the door clamp may be rotated backwards through 90 degrees into the neutral position and the airlock crank turned clockwise, opening the inner gate. During this time the shutter, whose upper surface is coated with fluo-

rescent material and serves as viewing screen, will normally be closed. The graduated head at the left of the table is now rotated clockwise, moving the plate into position under the shutter, the plate cover being held back by a stop. Each graduation on the head corresponds to

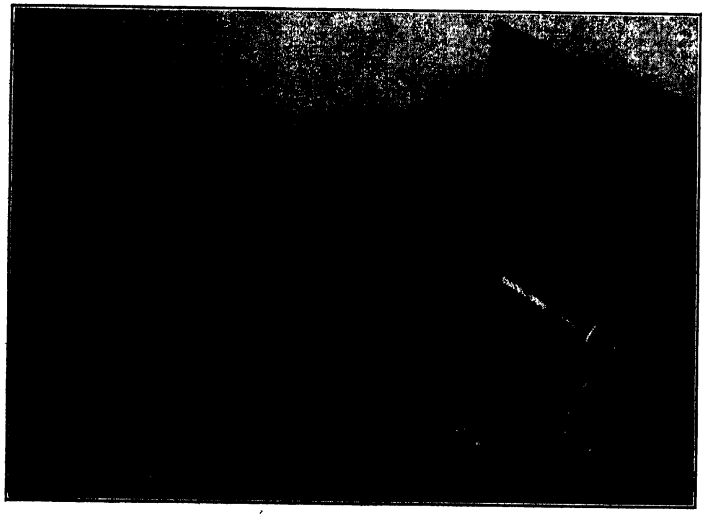


FIG. 5·28. Plate Chamber of the Type B Electron Microscope.

a displacement of the plate by ½ inch. The opening of the inner gate, acting as frame, is 3 by 2 inches. Thus three pictures of this size can be obtained on a single plate by rotating the head by six graduations between successive exposures. The exposure itself is made by raising the screen, accomplished by rotating the screen control head (Fig. 5·29) which retracts a bent rod engaging a bar attached to the screen. Here, also, the vacuum connection is maintained by a metal bellows. A similar arrangement, also visible in Fig. 5·29, serves to adjust a mask determining the picture width. This mask consists of a frame whose inclination can be varied, the portion of the beam which is transmitted depending on the angle of inclination. The use of this mask in conjunction with the graduated plate displacement permits the taking of eighteen ½-inch pictures, nine 1-inch pictures, or five 2-inch pictures, etc., on the same plate.

The principle of the Siemens plate lock is shown in Fig. 5·30. At the top, the inner gate plate (which is also coated with fluorescent material, serving as a second viewing screen) is pushed up against the opening

into the microscope by means of an eccentric cam and the plate (2½ by 3½ inches) is introduced through an open door. In the central figure the outer door has been closed and the plate chamber has been exhausted by the forepump. As a result, the inner gate plate drops down on the

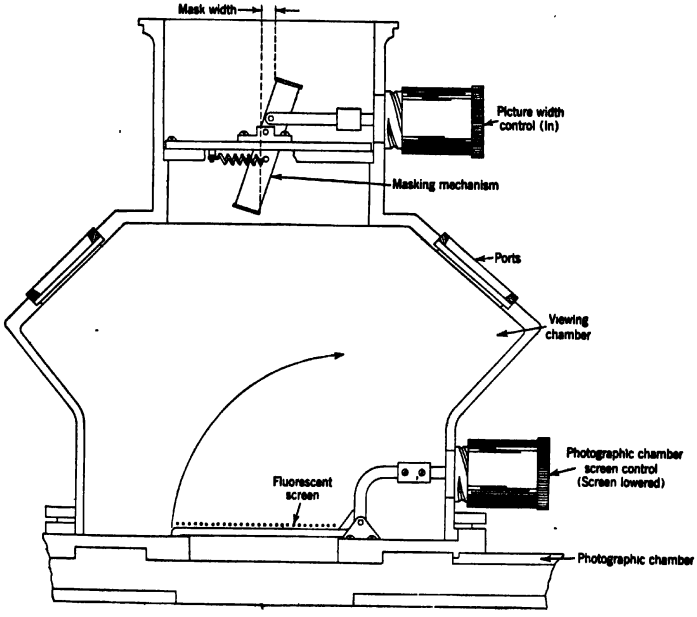


FIG. 5·29. Viewing Chamber of the Type B Electron Microscope (Section).

plate holder and engages a projection in its cover as the cam is turned back. Finally, after the shutter, acting at the same time as a fluorescent screen, has been closed, a pinion engaging the rack of the inner gate plate moves it, together with the cover, to the right. Now the plate is ready for exposure. As in the Type B electron microscope, as many operations as possible have been interconnected in this plate airlock, so as to simplify the procedure. Ground joints are employed in place of metal bellows for communicating external manipulations to the evacuated plate chamber. It is obvious that the arrangement here described does not lend itself to the taking of a series of pictures on the same plate.

The plate lock of the von Ardenne microscope is identical in principle with that of the Siemens instrument, with the distinction that the plate cover is retained in fixed position and the open plate holder, together with the inner gate plate, translated horizontally. This makes it

possible to take four pictures (approximately 1¾ by 1¼ inches in size) on a pair of 1¾ by 2½-inch Schumann plates placed side by side in the plate holder. No attempt has been made to simplify the manipulation by the combination of different operations as in the two previously described plate chambers.

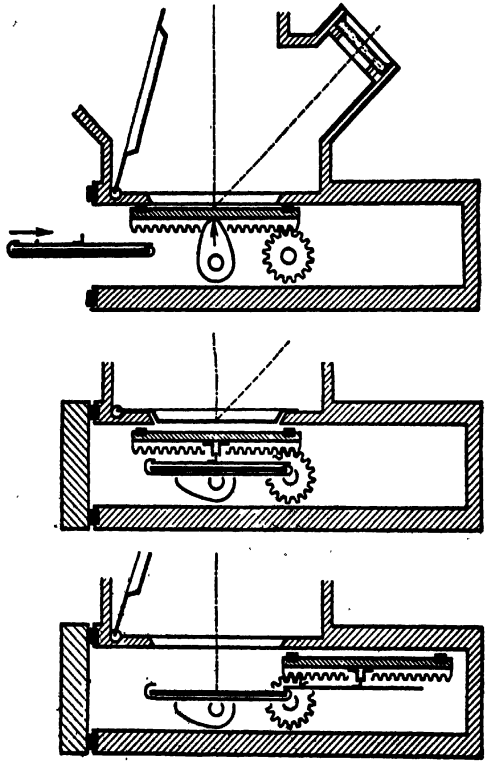


FIG. 5·30. Plate Lock of the Siemens Electron Microscope (Section). (v. Borries and Ruska, *Ergebn. d. exakt. Naturwiss.*, Vol. 19, pp. 237–322, 1940. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

In the small electron microscope, the plate holder, containing one 2 by 2-inch plate, is introduced into the bottom of the viewing chamber through a door below the fluorescent screen (Fig. 5·6). The counter-clockwise rotation of a knob on the right of the door raises the plate holder to a position in front of the screen and, if continued, locks it there in place, releasing, simultaneously, the plate cover (Fig. 5·8). A rota-

tion of the knob in a clockwise direction then drops the cover down, permitting the exposure to be made. Reversing the direction of rotation once more, the cover is replaced and the plate holder is unlatched. A final clockwise rotation returns the closed plate holder to its original position, from which it may be removed through the door below the screen.

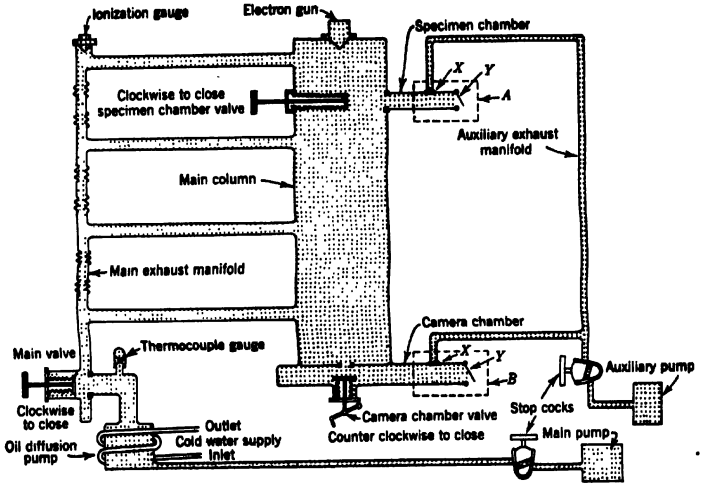


FIG. 5·31. Schematic Diagram of Vacuum System of Type B Electron Microscope.

5·6. The Vacuum System. The gas pressure within the microscope should, during operation, be low enough that (1) gas discharges do not take place between the cathode and anode, (2) the life of the cathode is not reduced appreciably by oxidation or positive-ion bombardment, and (3) collisions between the beam electrons and gas molecules are sufficiently rare that electrons scattered as a consequence do not appreciably reduce the contrast of the image. It is found that all these conditions are satisfied if the pressure is reduced to approximately 10^{-4} mm Hg. The mean free path of the electrons is then nearly 3 meters.

A schematic diagram of the complete vacuum system of the Type B electron microscope is shown in Fig. 5·31. It is seen to possess two separate pumping systems: the main pumping system for evacuating the body of the microscope and the auxiliary pumping system for exhausting the airlock chambers. This arrangement eliminates the possibility of a deterioration of the main vacuum during the pumping out of the chambers.

The main pumping system consists of a rotary oil pump, capable of reducing the pressure in a vessel from atmospheric pressure (760 mm Hg) to approximately 10^{-2} mm Hg, and an electrically heated oil diffusion pump, the two being connected by rubber pressure tubing. The forepump, together with the auxiliary pump for the evacuation of the airlock chambers, is set up separately in a soundproof housing at a small distance from the microscope so as to prevent the communication of mechanical vibrations. The diffusion pump is rigidly connected to the body of the microscope. The employment of oils with low vapor pressure (*octoil* or *apiezon* oils) in the diffusion pump in place of mercury makes this possible, since the boiling of the oil is not accompanied by "bumping." With a mercury diffusion pump a flexible connection between the pump and the microscope is required. A further advantage of the employment of oil is that a freezing-out trap is not required since at room temperature the vapor pressure of the oils employed ranges between 10^{-5} and 10^{-7} mm Hg, whereas that of mercury is about $3 \cdot 10^{-3}$ mm Hg.

The speed of evacuation at low pressures is proportional not only to the difference in pressure between the vessel to be evacuated and the diffusion pump intake, but also to the cube of the diameter and the reciprocal of the length of the connecting vacuum conduit. Since, in the electron microscope, the channel is severely constricted at several places — especially at the lenses — it is advantageous to have a number of pump connections from a large-diameter manifold. As shown in Fig. 5·3, such connections are made at the viewing chamber, at the intermediate tube, at the object chamber, and, finally, at the gun. The several junctions are connected by flexible metal bellows to permit the displacement of the sections of the microscope relative to each other during the alignment. A stopcock is placed between the two pumps to prevent forepump oil from entering the connecting tube when the forepump is shut off. In addition, a large valve is located between the diffusion pump intake and the exhaust manifold of the microscope (Fig. 5·32). This permits maintaining the diffusion pump at pumping temperature even when (for example, for exchanging the objective pole pieces) air is admitted into the body of the microscope.

The auxiliary system consists of a second rotary oil pump, connected directly, by flexible tubing and metal pipes, to the two airlock chambers. These chambers are at the same pressure as the main chamber of the microscope during operation since they communicate directly with it, the connection to the auxiliary system being closed off.

Two gauges measure the vacuum within the instrument. A thermocouple gauge is attached between the diffusion pump intake and the

main valve. It gives a continuous check on the operation of the diffusion pump. The gauge contains a thermocouple junction which is heated to a temperature of, at most, 300° C by a thin filament carrying about 100 milliamperes (Fig. 5·33*a*). The current generated in the thermocouple circuit is read on a galvanometer. Since, for constant heater current, a smaller proportion of the heat liberated within the

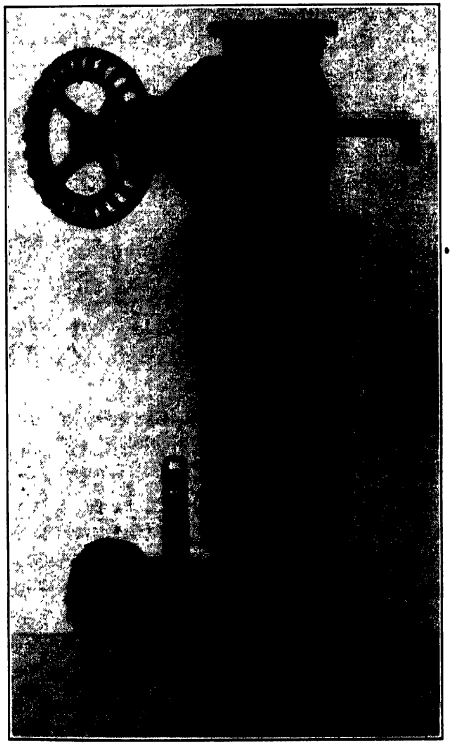


FIG. 5·32. Oil Diffusion Pump and Main Vacuum Valve of Type B Electron Microscope.

heater wire is carried off by colliding gas molecules at low pressures than at high pressures, the temperature of the junction, and, hence, the galvanometer current will increase as the pressure is reduced. The heater current and the galvanometer sensitivity are adjusted so that a full deflection of the galvanometer needle corresponds to a sufficiently good vacuum. The sensitivity of this device becomes very small by the time

a vacuum of 10^{-3} mm Hg has been attained, so that the thermocouple gauge is not applicable for measuring pressures materially lower than this.

In addition to the thermocouple gauge, a high-sensitivity ionization gauge is attached to the manifold near the electron gun, the vacuum conditions at this point being most important for the satisfactory operation of the microscope. The ionization gauge (Fig. 5·33*b*) is an

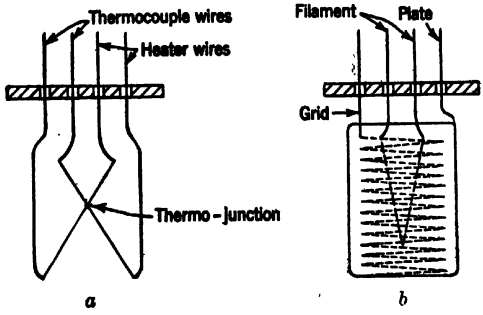


FIG. 5·33. Vacuum Gauges: (*a*) Thermocouple Gauge; (*b*) Ionization Gauge.

ordinary three-electrode vacuum tube with the grid made positive and the plate negative with respect to the filament cathode. Electrons leaving the cathode are accelerated toward the grid. In part, they are collected by the grid. A large proportion, however, flies through the grid into the retarding field in front of the plate, which causes the electrons to turn back towards the grid. These electrons, too, are finally collected by the grid, though frequently only after a large number of oscillations. While passing through the space between the plate and the grid they may collide with gas molecules and ionize them, the positive ions being then drawn to the plate. At low pressures the number of such collisions will be directly proportional to the gas pressure in the tube. Thus, if the grid current is adjusted to a fixed value (for example, 5 milliamperes), the plate current will be directly proportional to the pressure in the tube. In a particular type of ionization gauge employed in the Type B electron microscope a plate current of 10 microamperes corresponds to a pressure of $3.5 \cdot 10^{-4}$ mm Hg. Any reading lower than this may thus, here, be regarded as indicating a satisfactory vacuum.

As mentioned above, the Siemens and von Ardenne microscopes employ mercury diffusion pumps, heated by gas or electricity. This may constitute a reason for the desirability of establishing direct contact between the object mount and the objective so as to reduce the sensitiv-

ity to mechanical vibrations. Both instruments rely on a single pump connection to the microscope to exhaust the whole system. The forepump is also used to evacuate the plate chamber. As check on the vacuum in the forepump system the Siemens microscope employs a simple mercury manometer; as indicator for the main system, a small discharge tube, to which voltage may be applied by pressing a button. If no discharge appears, the vacuum is regarded as adequate for the operation of the microscope.

The distinguishing features of the vacuum system of the small electron microscope are visible in Fig. 5·8. They are, in particular, a single-jet oil diffusion pump having a speed of 10 liters per second and permitting the achievement of an ultimate vacuum of $5 \cdot 10^{-5}$ mm Hg and a valve mechanism which causes the whole pumping cycle to be traversed with the up-and-down motion of a single sliding member. The sliding member is a flat plate which, by means of cam surfaces and rods vacuum-sealed with packing glands, controls three individual valves. The valves are so interconnected that in the intermediate position the forepump is connected to the diffusion pump and the microscope chamber is sealed off from both. In this position only can air be admitted by actuating a push button. In the lowest position the diffusion pump is sealed off and the forepump connected with the microscope chamber. Finally, in the top position, the microscope chamber is connected to the diffusion pump and the diffusion pump to the forepump. Only in this operating position is it possible to turn on the high voltage and the filament current.

5·7. Controls and Adjustments. The procedure of putting an electron microscope into operation varies relatively little from instrument to instrument. It will thus be adequate to outline that followed with the Type B electron microscope. If it is assumed that the instrument has not been in use for some time, so that all the stopcocks and valves are closed and the switches are open, the first step is to start the rotary oil pumps. Next the stopcocks in the pump lines and, then, the main valve between the diffusion pump and the microscope are opened. Now the flow of water through the cooling jacket of the diffusion pump is started and the diffusion pump heater is turned on. Any failure in the flow of water automatically cuts off the heater current, which in turn is indicated by a colored light on the control panel. The thermocouple galvanometer reading is now watched, the thermocouple heater current having been adjusted to the proper value. As the reading approaches full scale, the ionization gauge is turned on. When the plate current has dropped to an adequately low value (for example, 10 microamperes), the main power switch is closed. After about a minute of permitting the tubes of

the power supplies to warm up (the process may be followed with the aid of a test meter mounted on the control panel), the high voltage, with the high-voltage selector adjusted to the desired point, is ready to be turned on and the switches for the lens currents closed. Overloads due, for example, to gas discharges in the microscope cause a relay to interrupt the high-voltage circuit and to light a neon lamp on the panel. A push button serves to reset the relay as soon as the cause of the overload is removed. With the voltage, lens currents, and filament current turned on the microscope is, finally, ready for operation.



FIG. 5·34. 300-Kilovolt Electron Microscope (RCA Laboratories). (Courtesy *J. Applied Phys.*, reference 20.)

5·8. Modifications for High-Voltage Operation. It has been pointed out before[17] that the use of very high-speed electrons, or large accelerating voltages, is required if sub-light-microscopic detail is to be distinguished in objects of the order of a micron in thickness. They become essential, in particular, in the study of biological sections, which can scarcely be prepared thinner than this. It is, hence, not surprising that work looking toward an enhancement of the operating voltage of the microscope has gone on in several laboratories. Thus von Ardenn [18]

[17] See section 4·7.
[18] See reference 18.

gives a description of a 200-kilovolt Universal Electron Microscope. Similarly, Müller and Ruska[19] report results obtained with a 220-kilovolt installation utilizing a modified Siemens microscope. An experimental high-voltage microscope designed for a maximum voltage of 300 kilovolts, constructed in the RCA Laboratories,[20] is shown in Fig. 5·34, the high-voltage and x-ray protection having been removed to show more clearly the several parts of the assembly.

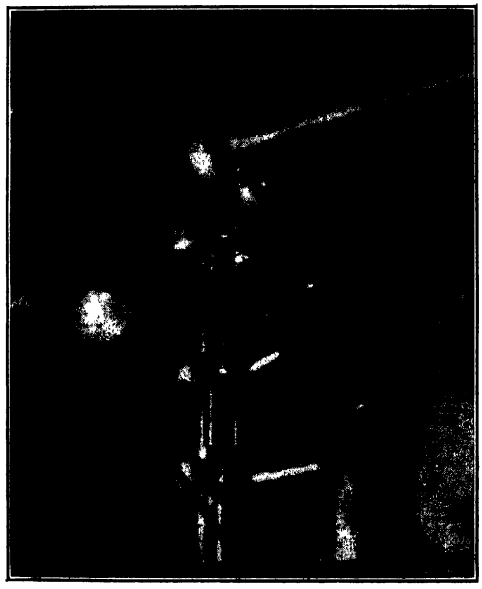


FIG. 5·35. Three-Section Electron Gun for the High-Voltage Electron Microscope. (Courtesy *J. Applied Phys.*, reference 20.)

The microscope column is the same as that used in the Type B electron microscope with the exception of the electron gun. Owing to the much greater volume of the high-voltage generator, it was necessary to place it in a separate large oil tank, connection of the high-voltage terminal to the cathode being made by a polished aluminum tube. The driving circuits in the microscope cabinet, in turn, are connected to the circuits within the tank by concentric cables and spark-plug connectors passing through the tank wall.

The gun shown in Fig. 5·34 is a two-stage gun, permitting stable

[19] See reference 19.

[20] See Zworykin, Hillier, and Vance, reference 20.

operation up to 200 kilovolts. The voltage is distributed between the stages by connecting a high-voltage divider across them. In order to utilize the full available voltage (300 kilovolts), it became necessary to add one more stage, resulting in the gun shown in Fig. 5·35. The intermediate electrodes were shaped so that they would collect a minimum of electron current. Since, with the new, high, gun structure, the instrument became quite sensitive to mechanical vibrations, a special brace had to be provided between the cabinet in the rear and the inter-



FIG. 5·36. A Group of *B. megatherium* Taken with 50- and 200-Kilovolt Electrons. Magnification: 6500.

mediate tube and, in addition, both the microscope and the cabinet had to be placed on thick pads of sponge rubber. In addition, a large amount of lead shielding had to be provided, especially about the gun, to reduce the x-ray intensity around the instrument (at some points reaching a value of 10^{-3} *r* per second) to a safe amount. The total installation weighs several tons, and required the reinforcement of the laboratory floor.

Figure 5·36 shows pictures of the same specimen of *Bacillus megatherium* taken with 50-kilovolt and 200-kilovolt electrons. It is seen that, although considerable internal structure is brought out in the 200-kilovolt picture, none can be distinguished in the low-voltage micrograph. This is found to be the case even if the micrograph is printed very lightly.

However, in the high-voltage picture the contrast of some low-density detail outside the bacteria has been reduced to such an extent as to render it invisible.

The experience with this instrument indicates that, as might have been expected, the only advantage gained from the employment of higher voltages rests in the greater penetration of thick specimens. Furthermore, this advantage becomes significant only in passing from the usual operating voltages of the commercial instruments (about 60 kilovolts) to 150 kilovolts or beyond. Resolutions of at least 100 A.U. were obtained at 200 and 250 kilovolts. Magnifications up to 10,000 were achieved at 250 kilovolts by employing a special projector lens.

REFERENCES

1. B. v. Borries and E. Ruska, "A supermicroscope for research laboratories," *Naturwissenschaften*, Vol. 27, pp. 577–582, August 1939.
2. B. v. Borries and E. Ruska, "The technique of the Siemens Supermicroscope," *Siemens-Z.*, Vol. 20, pp. 217–227, November–December 1940.
3. V. K. Zworykin, J. Hillier, and A. W. Vance, "An electron microscope for practical laboratory service," *Elec. Eng.*, Vol. 60, pp. 157–161, April 1941.
4. J. Hillier and A. W. Vance, "Recent developments in the electron microscope," *Proc. Inst. Radio Engrs.*, Vol. 29, pp. 167–176, April 1941.
5. M. v. Ardenne, "A universal electron microscope for bright-field, dark-field, and stereo operation," *Z. Physik*, Vol. 115, pp. 339–368, March 1940.
6. V. K. Zworykin and J. Hillier, "A compact high resolving power electron microscope," *J. Applied Phys.*, Vol. 14, pp. 658–673, December 1943.
7. E. Ruska, "On progress in the construction and performance of the magnetic electron microscope," *Z. Physik*, Vol. 87, pp. 580–602, February 1934.
8. L. Marton, "Electron microscopy of biological objects, III," *Bull. acad. roy. Belg.*, Classe des Sciences, Vol. 21, pp. 606–617, 1935.
9. L. C. Martin, R. V. Whelpton, and D. H. Parnum, "A new electron microscope," *J. Sci. Instruments*, Vol. 14, pp. 14–24, January 1937.
10. E. Driest and H. O. Müller, "Some electron micrographs," *Z. wiss. Mikroskop.*, Vol. 52, p. 53, 1935.
11. A. Prebus and J. Hillier, "The construction of a magnetic electron microscope of high resolving power," *Can. J. Research*, Vol. A17, pp. 49–63, April 1939.
12. E. F. Burton, J. Hillier, and A. Prebus, "A report on the development of the electron supermicroscope at Toronto," *Phys. Rev.*, Vol. 56, pp. 1171–1172, November 1939.
13. C. E. Hall and A. L. Schoen, "Application of the electron microscope to the study of photographic phenomena," *J. Optical Soc. Am.*, Vol. 31, pp. 281–285, April 1941.
14. L. Marton, M. C. Banca, and J. F. Bender, "A new electron microscope," *RCA Review*, Vol. 5, pp. 232–242, October 1940.
15. L. Marton, "A new electron microscope," *Phys. Rev.*, Vol. 58, pp. 57–60, July 1940.
16. E. Kinder and A. Pendzich, "A new magnetic lens of short focal length," *Jahrb. der AEG-Forschung*, Vol. 7, pp. 23–26, March 1940.

17. B. v. Borries, E. Ruska, J. Krumm, and H. O. Müller, "Supermicroscopic imaging by magnetostatic lenses," *Naturwissenschaften*, Vol. 28, pp. 350–351, May 1940.
18. M. v. Ardenne, "On a 200-kilovolt universal electron microscope with object-shading arrangement," *Z. Physik*, Vol. 117, pp. 657 ff., 1941.
19. H. O. Müller and E. Ruska, "A supermicroscope for 220 kilovolt beam voltage," *Kolloid-Z.*, Vol. 95, pp. 21–25, March 1941.
20. V. K. Zworykin, J. Hillier, and A. W. Vance, "A preliminary report on the development of a 300-kilovolt magnetic electron microscope," *J. Applied Phys.*, Vol. 12, pp. 738–742, October 1941.

CHAPTER 6

ABERRATIONS AND TOLERANCES IN THE ELECTRON MICROSCOPE

6·1. Spherical Aberration and Diffraction. The two most important fundamental limitations to the resolution obtained with electron microscopes are those resulting from diffraction and spherical aberration. The limitation of the resolution of the instrument due to the first is proportional to the wave length λ of the electrons and to the reciprocal of the effective aperture α of the objective:

$$d_{\text{diff}} = \frac{0.61\lambda}{\alpha} \qquad [6\cdot1]$$

The spherical aberration of the objective, on the other hand, limits the resolution to:[1]

$$d_{\text{sph}} = 0.4Cf\alpha^3 \qquad [6\cdot2]$$

where C is a dimensionless constant characteristic of the objective and f the focal length of the latter. Values of Cf as low as 0.2 centimeter are obtained in modern instruments.[2]

If the sum of these two defects is minimized by choosing the most favorable value of the aperture α, the limitation of the resolution introduced by these two defects is given, according to Eq. 19·50*b*, by

$$d_{\text{min}} = 0.95\lambda^{3/4}(Cf)^{1/4}$$

Figure 6·1 shows d_{min}, as well as the corresponding values of the optimum effective aperture α_{opt}, as functions of Cf.

With perfect alignment, shielding, and current and voltage stability and electrons of uniform initial velocity, these resolving powers may be approached under either of two circumstances: (1) with arbitrary conditions of illumination, an object too thin to cause appreciable velocity losses on the part of a large fraction of the electrons, and a physical objective aperture just large enough to admit imaging pencils of the aperture α_{opt}; and (2) with parallel illumination (illuminating aperture much smaller than α_{opt}), an object too thin to cause a large amount of scattering out of the cone with half-vertex angle α_{opt}, and either no

[1] See section 19·3.

[2] See, e.g., v. Borries and Ruska, reference 1, and Dosse, reference 2.

physical aperture in the objective or one large compared to that required for the first case. If the illuminating angle (and physical aperture) is large compared to α_{opt}, the resolution will be given simply by Eq. 6·2; on the other hand, if the physical aperture of the objective is made small compared to that which will just accept pencils of aperture α_{opt}, the resolution will be determined exclusively by the diffraction effect, that is, by Eq. 6·1. The scattering through large angles, which takes place in thick specimens, has an effect similar to that of increasing the illuminating aperture.

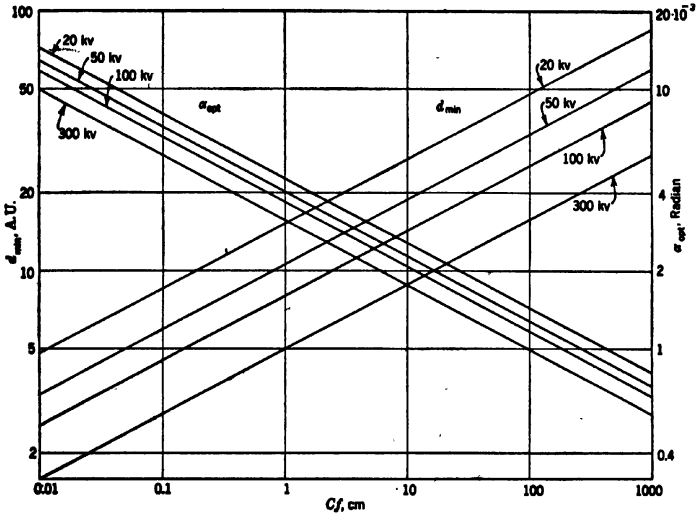


FIG. 6·1. Variation of the Least-Resolvable Separation d_{min} and the Corresponding Optimum Objective Aperture α_{opt} as Function of the Spherical-Aberration Coefficient of the Objective Cf.

6·2. Chromatic Aberration. It is brought out in section 16·6 that, if the range of variation of the kinetic energy eV of the imaging electron is $e\Delta V$, the diameter of the disk of confusion due to axial chromatic aberration is

$$d_{chr} = K\left(\frac{\Delta V}{V}\right)\cdot f\cdot\alpha \qquad [6\cdot3]$$

where K is a constant which, for magnetic lenses, is slightly less than unity and f is the focal length of the lens. If the representation in Fig. 17·25 is utilized, it is reasonable to put $K = 0.75$ for the magnetic objectives employed in electron microscopes. The same figure suggests

that for electrostatic objectives with the center electrode connected to the cathode K is about five times as large as this.

As previously mentioned, the three sources of velocity variations in the electron microscope are variations in the initial energy of emission of the electrons, velocity losses within the object, and the fluctuations of the accelerating voltage.

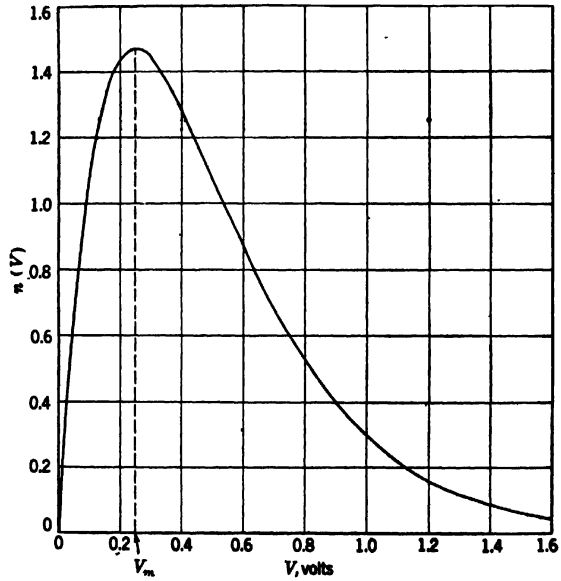


FIG. 6·2. Maxwellian Distribution of the Electrons Emitted by a Thermionic Cathode.

The initial velocities of the electrons leaving the thermionic cathode are distributed according to the Maxwellian law. If $n(V)\cdot dV$ is the fraction of electrons emitted with an initial kinetic energy between eV and $e(V + dV)$,

$$n(V)\cdot dV = \left(\frac{e}{kT}\right)^2 \cdot \exp\left(-\frac{eV}{kT}\right) \cdot V \cdot dV \qquad [6\cdot4]$$

where k is Boltzmann's constant ($1.38\cdot10^{-16}$ erg per degree) and T is the absolute temperature of the cathode. This distribution (Fig. 6·2) has a maximum at

$$V_m = \frac{kT}{e} = \frac{T}{11{,}600} \text{ volt}$$

Thus, for an operating temperature of $T = 2900°$ A, suitable for a tungsten cathode, $V_m = 0.25$ volt. For $K = 0.75$, $V = 50{,}000$ volts, $f = 0.4$ centimeter, and $\alpha = 3 \cdot 10^{-3}$ radian, this leads, with ΔV equal to V_m, to

$$d_{\text{chr}} = 0.75 \cdot 10^{-5} \cdot 0.2 \cdot 3 \cdot 10^{-3} = 5 \cdot 10^{-9} \text{ cm (0.5 A.U.)}$$

With a thermionic cathode the initial velocities of the electrons thus do not appreciably affect the sharpness of the image.

The effect of the object thickness (or, more strictly, object mass thickness) on the velocity distribution of the electrons and, hence, on the sharpness of the image has already been discussed in section 4·7. It can be minimized by favorable techniques of object preparation and the use of a very thin, or no, supporting film as well as by choosing physical object apertures of optimum dimension. Beyond this point it is a basic property of the specimen.

The final source of chromatic aberration, fluctuations in the voltage supplies, may be controlled by properly stabilizing these sources of power. The degree of stabilization required to keep the consequent aberrations below a level which will appreciably affect the performance of the microscope will be discussed in section 6·4.

6·3. Field Aberrations. The aberrations discussed above affect all points of the image equally. They are introduced almost entirely by the objective. If the objective and projector have equal magnification and are otherwise similar, the projector contributes an amount less than the objective by a factor equal to the first or second power of the total magnification of the instrument. However, as brought out in section 1·6 and discussed in greater detail in sections 16·2 and 16·6, there are a number of other aberrations — coma, curvature of field and astigmatism, distortion, and chromatic difference in magnification and in rotation — which vanish on the axis and become increasingly prominent in proceeding from the center of the image outward. These *field aberrations* are due to both the projector and the objective, the projector playing the predominant role in distortion, the objective in coma. For the other field aberrations the contributions of the objective and projector are of the same order. In light optics this fact is utilized to correct chromatic difference in magnification, curvature of field, and astigmatism introduced by the objective with the aid of the eyepiece (for example, in *compensating eyepieces*). A similar balancing of spherical aberrations, axial chromatic aberrations, and coma (*sine condition*) of the two lens systems is not practicable, since their effects on the image are of two different orders of magnitude. These defects must be corrected in the

objective. The same holds for distortion, which must be corrected in the eyepiece or projector.

Among the geometric (as distinguished from the chromatic) field aberrations, distortion is the most important. It results in an apparent displacement of the points of the object, as judged from the image, by an amount[3]

$$d_{\text{dist}} = K_1 M_1^2 r_o^3 \qquad [6\cdot5]$$

Here r_o is the distance of the object point considered from the optic axis of the objective and M_1 is the objective magnification. The displacement d_{dist} may be in any direction. At times it is nearly tangential, the sense of the tangential component reversing with the sense of the current in the projector coil. On this *rotational* or *anisotropic* distortion is superposed an outward radial component, that is, " pincushion " distortion. A typical image obtained with excessive distortion of this kind is shown in Fig. 6·3. For weak lenses the rotational distortion recedes relative to the pincushion distortion. In the limit the distortion is pure pincushion distortion, and the factor K_1 is given by C/f^2, where C is the dimensionless spherical-aberration constant of the projector and f its focal length.

Whereas distortion merely displaces the apparent position of any object point, curvature of field and astigmatism diffuse it into an elliptical disk. Its major axis, which is proportional to the square of the distance of the object point from the axis and to the first power of the objective aperture, may be expressed by[4]

$$d_{ca} = 2K_2 \cdot r_o^2 \cdot \alpha \qquad [6\cdot6]$$

The insertion of reasonable values for the several quantities in the equation indicates that curvature of field and astigmatism are not at all serious sources of unsharpness in the electron microscope. For example, with $K_2 = 15$ centimeter^{-1}, $r_o = 2\cdot10^{-4}$ centimeter (diameter of image = 4 centimeters), $M = 10{,}000$, and $\alpha = 3\cdot10^{-3}$ radian, $d_{ca} =$

[3] $K_1 = (S_1^2/M_2^2 + S_2^2/M_2^2)^{1/2}$, where M_2 is the magnification of the projector and S_1 and S_2 are its coefficients of isotropic and anisotropic distortion, respectively (see section 16·2). The angle of the displacement with the radial direction is given by arctan (S_2/S_1). In the example discussed in section 17·5, $K_1 = 6.9$ (pole-piece diameter)$^{-2}$ and the angle is 71 degrees.

[4] $$K_2 = f_1\left\{\left(\frac{S_{31}}{M_1} + \frac{S_{32}f_2}{M_2 f_1}\right) + \left[\left(\frac{S_{41}}{M_1} + \frac{S_{42}f_2}{M_2 f_1}\right)^2 + \left(\frac{S_{51}}{M_1} + \frac{S_{52}f_2}{M_2 f_1}\right)^2\right]^{1/2}\right\}$$

For example, for $M = 10{,}000$, $f_1 = 0.5$ cm, the optical system described in section 17·4 yields $K_2 = 15$ cm^{-1}. The subscript 1 refers to the objective, the subscript 2, to the projector. S_3, S_4, and S_5 are the aberration coefficients of curvature, astigmatism, and anisotropic astigmatism, respectively.

$3.6 \cdot 10^{-9}$ centimeter (0.36 A.U.). The orientation of the ellipse is arbitrary.

The coma of the objective also has very little effect on the quality of the electron-microscope image. It spreads an object point out into an aberration figure, whose maximum diameter is given by[5]

$$d_{\text{coma}} = K_3 \cdot r_o \cdot \alpha^2 \qquad [6\cdot7]$$

The orientation of the aberration figure, again, varies from lens to lens. With $K_3 = 3.5$, $r_o = 2 \cdot 10^{-4}$ centimeter, and $\alpha = 3 \cdot 10^{-3}$ radian, $d_{\text{coma}} = 6 \cdot 10^{-9}$ centimeter (0.6 A.U.).

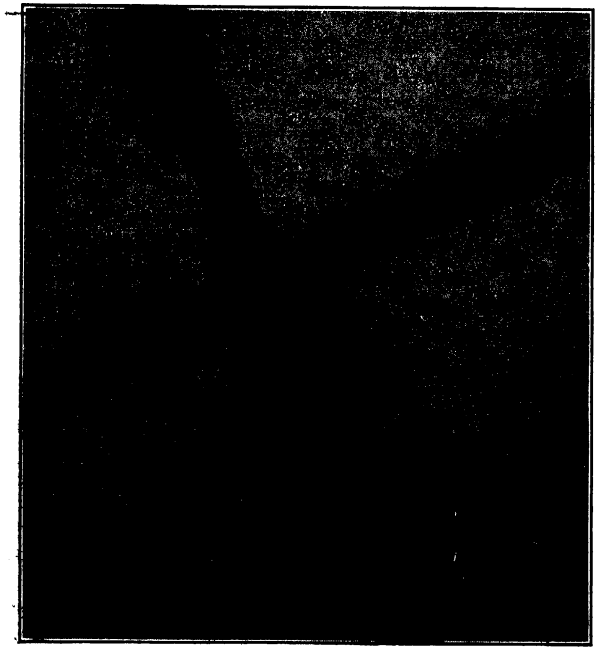


FIG. 6-3. Electron Micrograph of a Uniform Taut Wire Showing Excessive Pincushion and Rotational Distortion (0.001-Inch Tungsten Wires).

The chromatic field aberration, compounded of chromatic difference in magnification and chromatic difference in rotation, draws each object point out into a line whose length is proportional to the range of kinetic energies $e\Delta V$ represented in the imaging pencils and to the distance of

[5] $K_3 = 3\beta_1^2(S_0^2/M_1^2 + S_1^2/M_1^2)^{1/2}$. In the example cited above, $K_3 = 3.5$.

the point from the axis:[6]

$$d_{\text{chr.f.}} = K_4 \cdot \frac{\Delta V}{V} \cdot r_o \qquad [6\cdot8]$$

For $K_4 = 1$, $\Delta V/V = 10^{-4}$, $r_o = 2\cdot10^{-4}$, $d_{\text{chr.f.}} = 2\cdot10^{-8}$ centimeter (2 A.U.). It is seen that this source of unsharpness is more serious than curvature, astigmatism, and coma. The radial component of the displacement is due to chromatic difference of magnification. As long as the object is outside the lens field, this component is given by

$$d_{\text{d.m.}} = \frac{\dfrac{dM}{dV}}{\dfrac{M}{V}} \cdot \frac{\Delta V}{V} \cdot r_o \qquad [6\cdot9]$$

This equation, furthermore, applies quite generally for the projector. For weak lenses, the coefficient $(dM/dV)/(M/V) = -1$.

For the objective, the coefficient in the expression for $d_{\text{d.m.}}$ first decreases numerically as the lens strength is increased, passes through zero, and then, though the magnification of the objective increases still further, changes sign and becomes increasingly positive. This behavior is made plausible by Fig. 6·4, *a* and *b*. In Fig. 6·4*a* the case of the object lying outside the lens field is illustrated. Increasing the voltage V causes the effective object plane (plane conjugate to the fixed image plane) to move backward from O to O'. Since the principal rays are parallel to the axis, the size of the effective object is unaltered, so that the image magnification is reduced by the same amount as though the object were moved physically back to O'. In Fig. 6·4*b*, on the other hand, the object lies well within the lens field. As the voltage is increased, the effective object plane again moves back to O'. Now, however, because of the inclination of the principal rays, the size of the object pattern in O' is larger than the true object and the image magnification correspondingly larger than that obtained with a refocusing of the actual object in the plane O' at the higher voltage. The increase in size due to the inclination of the principal rays may easily outweigh the reduction

[6] In terms of the chromatic-aberration coefficients C_1 and C_2 (section 16·6):

$$K_4 = \frac{V}{M\Delta V}\left[(C_{11} + C_{12})^2 + (C_{21} + C_{22})^2\right]^{1/4}$$

The second indices refer to the objective (1) and the projector (2), respectively. The example here discussed (see section 17·6) leads to the value $K_4 = 2.43$ with an angle of inclination of 68 degrees if the current flow has the same sense in the two lens coils. If the sense of flow is opposite, $K_4 = 0.93$, and the angle of inclination to the radial direction is 11 degrees.

in lens magnification (with the object in perfect focus) at the higher voltage. For instance, in the special case of the strong magnetic objective discussed in section 17·6, the coefficient in the expression for $d_{\text{d.m.}}$ is $+1.1$ as compared with the -1 applying for weak lenses. It is thus possible, for a particular value of the objective and the projector currents, to cause the chromatic difference in magnification to vanish.

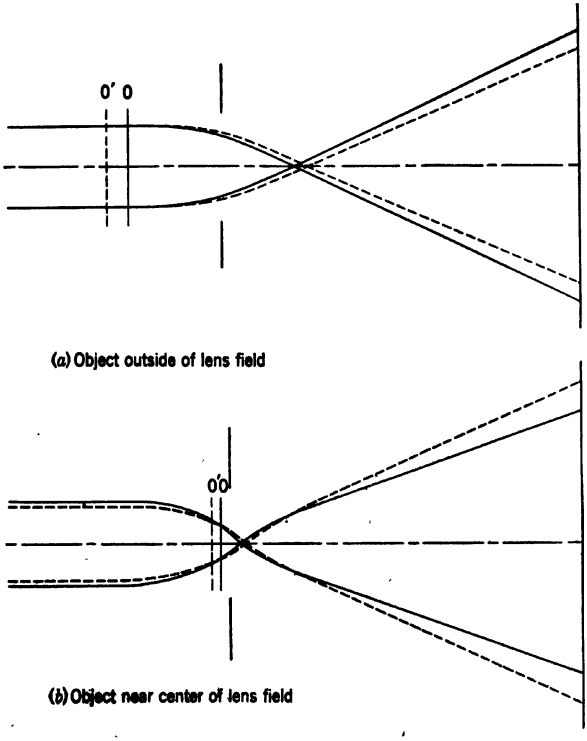


FIG. 6·4. Variation of Objective Magnification with the Kinetic Energy of the Electrons (Object in Fixed Position).

The chromatic difference in rotation is universally given by

$$d_{\text{d.r.}} = \frac{\chi}{2} \cdot \frac{\Delta V}{V} \cdot r_o \qquad [6\cdot10]$$

where χ is the angle of image rotation produced by the lens in question (Eq. 16·103*b*). It represents the tangential component of the chromatic field aberration. Upper limits for the coefficient are here readily

established. For the objective (if no intermediate images are formed) χ cannot exceed π (3.142); for the projector it cannot exceed 2π (6.283). If account is taken of the fact that the projector will always be operated with coil currents less than that required to yield the maximum magnification of which it is capable (Fig. 4·19), the limitation of χ for the projector will correspond closely to that for the objective. Here, again, as the value of χ reverses sign with the direction of the coil current, the opportunity exists to compensate the chromatic difference in rotation introduced by the objective by that introduced by the projector by making the sense of current flow in the two lens coils opposite. The exact vanishing of both the chromatic difference of magnification and that of rotation at the same point — and hence the vanishing of K_4 in Eq. 6·8 — is made possible only by controlling the field distributions as well as the field strengths in the coils. The relation between the total chromatic field aberration and the chromatic differences of magnification and rotation is, of course,

$$d_{\text{chr.f.}} = (d_{\text{d.m.}}^2 + d_{\text{d.r.}}^2)^{1/2} \qquad [6\cdot11]$$

6·4. Current and Voltage Stabilization Tolerances. Whereas the geometric aberrations depend only on the character of the lens fields and, to a limited degree, on the placement and size of eventual limiting apertures and on the conditions of illumination, the chromatic aberrations are influenced as well by the unsteadiness of the high-voltage supply. Furthermore, variations in the currents within the lens coils have been shown to have an effect similar to that of unsteadiness of the high voltage. In fact, as long as the iron of the pole pieces and the casings is unsaturated, a given percentage fluctuation of the coil current has just twice the effect of an equal percentage fluctuation of the applied high voltage. Thus in this range the diameters of the circles of confusion on the axis due to voltage and current fluctuations will be

$$d_V = 0.75 f_1 \left(\frac{\Delta V}{V}\right) \alpha_1 \qquad [6\cdot12]$$

$$d_{I_1} = 1.5 f_1 \left(\frac{\Delta I_1}{I_1}\right) \alpha_1 \qquad [6\cdot13]$$

Here the subscript 1 indicates that the quantity in question pertains to the objective; the subscript 2, similarly, is used to indicate quantities pertaining to the projector. To include the effect of the projector on the circle of confusion on the axis, it would be necessary to add a similar term (with subscripts 2 in place of 1), divided by M_1^2, to the above two expressions. The only effect that saturation of the lens pole pieces can

have on Eqs. 6·12 and 6·13 consists of a reduction in the numerical coefficient (1.5) of the second equation. Thus the tolerance arrived at on the basis of Eqs. 6·12 and 6·13 may be regarded as conservative.

It may be stated in advance that the tolerances imposed by the axial chromatic aberration on the high voltage and on the objective current are sufficiently narrow so that no additional restrictions on them normally result from the requirement that the field aberrations should remain below a certain level. Field-aberration considerations fix, however, the permissible variations in the projector coil current. In view of Eqs. 6·8, 6·9, and 6·10, it is possible to write

$$d_{I_3} = 2K_4\left(\frac{\Delta I_2}{I_2}\right)r_o < (\pi^2 + 2^2)^{1/2}\left(\frac{\Delta I_2}{I_2}\right)r_o < 3.72\left(\frac{\Delta I_2}{I_2}\right)r_o \qquad [6\cdot14]$$

The upper limit given here again represents a very conservative estimate.

Equations 6·12, 6·13, and 6·14 show that the percentage tolerances in the applied voltage and in the objective current depend, apart from the permissible limitation of the resolution, only on the aperture and on the focal length of the objective, being inversely proportional to the latter. A reduction in objective focal length results in a proportionate increase in the permissible fluctuations. The percentage tolerance of the projector coil, however, depends only on the radius of the object area that is to be imaged with a variation in sharpness less than some prescribed amount.

Assume the objective focal length to be 0.4 centimeter, the aperture to be $3\cdot10^{-3}$ radian, and the area to be imaged with the prescribed sharpness to be 4 microns in diameter (image area at 10,000 magnification: 4 centimeters in diameter). Let the permissible unsharpness introduced by each source of aberration be at most 10 A.U. Then from the above three equations:

$$\frac{\Delta V}{V} = 1.1\cdot10^{-4} \qquad [6\cdot15a]$$

$$\frac{\Delta I_1}{I_1} = 0.55\cdot10^{-4} \qquad [6\cdot15b]$$

$$\frac{\Delta I_2}{I_2} = 1.3\cdot10^{-4} \qquad [6\cdot15c]$$

The last two tolerances — in particular that for the projector current

(6·15c) — may be regarded as extremely conservative. Furthermore, since fluctuations in the projector current introduce only an unsharpness in the marginal portions of the image (this being directly proportional to the current tolerance), fluctuations larger than that given by Eq. 6·15c by a factor as large as 5 do not seriously detract from the performance of the instrument.

Fluctuations in the condenser coil current affect the sharpness of the image only if no optimum physical aperture is employed in the objective and, even then, to a material degree only if the illuminating system is not well aligned. They may result in observable variations in the brightness of the image if the condenser is operated on the steep parts of the intensity curves in Fig. 4·6. To eliminate these inconveniences a certain amount of filtering in the condenser current supply is desirable if it is derived from an alternating-current source. A stability of 1 in 1000 is in all cases sufficient.

Tolerances may also be established in the case of the electrostatic microscope. If, here, the center electrodes of the objective and projector are connected directly to the cathode, the permissible fluctuations in the operating potential are determined by the relativistic aberration of the objective:[7]

$$d_{\text{rel}} = K_r \cdot \Delta V \cdot f_1 \cdot \alpha \qquad [6\cdot16]$$

where K_r has a value of $6.55\cdot10^{-7}$ volt^{-1} for weak lenses, decreasing to $2.47\cdot10^{-7}$ volt^{-1}, for a favorable design, as the lens strength is increased by reducing the center-electrode potential to that of the cathode. If the same values are assumed for the permissible unsharpness (10 A.U.), for the objective focal length (0.4 centimeter), and the aperture ($3\cdot10^{-3}$ radian) as before, the permissible voltage fluctuation becomes $\Delta V = 340$ volts or, with $V = 50{,}000$ volts,

$$\frac{\Delta V}{V} = 7\cdot10^{-3} \qquad [6\cdot17]$$

If the potentials of the center electrodes are made variable, for focusing and varying the magnification of the instrument, the voltage ratios of the electrodes of the objective must be kept constant within about one-fifth the percentage tolerance for the applied voltage in the magnetic microscope since the chromatic aberration of the electrostatic lenses is greater by a factor of the order of 5 than that of the magnetic objectives. The percentage tolerances for the projector voltage ratio and the projector current in the magnetic microscope become approximately equal.

[7] See section 18·2.

6·5. Construction and Alignment Tolerances. Since the accuracy of alignment and construction obviously has a great effect on the quality of the images obtained, it is of interest to estimate the maximum deviations from perfect symmetry which will still lead to a satisfactory image.

Consider, first, the horizontal centering of the objective with respect to the intermediate image aperture. A deviation from center here signifies simply a utilization of the outer parts of the image field of the objective. An examination of the several field aberrations of the objective indicates that, among them, curvature of field and astigmatism and the chromatic field aberrations are most serious. With $d_{ca} = 10^{-7}$ centimeter in Eq. 6·6 and with the same objective-projector combination as throughout the numerical examples in section 6·3, the permissible displacement of the objective relative to the intermediate image aperture becomes, for an objective magnification of 100, equal to 0.9 millimeter — if the voltage and objective-current stabilization is much superior to that prescribed by Eq. 6·15. For the same stabilization, the permissible displacement is only about one-tenth as large. The centering of the intermediate aperture relative to the projector should, similarly, be precise within a quantity of the order of 0.1 millimeter if no appreciable deterioration of the image is to be visible near its margin.

A tilt of the objective with respect to the axis of the instrument has two effects. On the one hand, it causes the central point to be imaged with field aberrations; on the other hand, it causes marginal points in the plane normal to the axis of tilt to lie in front of or back of the plane of focus by a distance $r_o\delta$, where δ is the angle of tilt. In view of the very small aperture of the electron-microscope objective and its consequent great depth of field (equal to $2{\cdot}10^{-7}/(3{\cdot}10^{-3}) = 7{\cdot}10^{-5}$ centimeter for a resolution of 10 A.U.) as well as its small object field, the second effect is unimportant. The effect of the field aberrations may be estimated if it is assumed (inaccurately) that they are equal to those for an object point a distance $f\delta$ from the axis of an untilted objective. Then, if curvature and astigmatism alone are considered, the permissible value of δ becomes $0.9{\cdot}10^{-3}/0.4 = 2{\cdot}10^{-3}$ radian (0.1 degree) for the same objective constants as before. Chromatic field aberrations will, again, require even more exact angular alignment, unless voltages and coil currents are held constant with extreme precision.

As far as image sharpness is concerned, the angular alignment of the projector is less critical than that of the objective by a factor of the order of the square root of the objective magnification. A corresponding inclination of 1 degree may, however, cause an excessive amount of image distortion, making more accurate angular alignment desirable.

Greatest precision is needed, of course, in the shaping and the relative

placement of the pole-piece apertures; or, in the case of electrostatic lenses, of the electrode apertures. However, it is very difficult to determine or even to state the requisite tolerances quantitatively. Since the magnetic field near the axis is, at any one point, a function of the complete pole-piece contours, very small irregularities in the contours have relatively small effect, since they tend to average out. In general, the greatest precision in machining is required for the surfaces facing both the gap and the optic axis. Receding from this central region the shaping becomes rapidly less important, especially if the iron is far from saturation. If difficulty is experienced in obtaining adequately perfect alignment of the two apertures, one may be made much larger than the other. In this case the lens field is relatively insensitive to the centering of the large aperture relative to the small one — an advantage achieved, it is true, at the expense of increased spherical aberration and reduced efficiency.

6·6. Shielding Requirements. The disturbing effect of a uniform magnetic field acting on the electron beam is discussed in detail in section 16·10. It is shown here that only the transverse component of this field is capable of seriously influencing the electron image. The apparent loss of definition produced by the action of the field in four different sections of the ray paths (see Fig. 4·27) is marked down separately below:

1. For the object-objective distance (length f_1):

$$d_1 = 0.148H\frac{f_1^2}{V^{1/2}} \qquad [6\cdot18a]$$

2. For the objective-intermediate image distance (length M_1f_1):

$$d_2 = 0.148HM_1\frac{f_1^2}{V^{1/2}} \qquad [6\cdot18b]$$

3. For the intermediate image-projector distance (length f_2):

$$d_3 = 0.148\frac{H}{M_1}\frac{f_2^2}{V^{1/2}} \qquad [6\cdot18c]$$

4. For the projector-final image distance (length M_2f_2):

$$d_4 = 0.148H\frac{M_2}{M_1}\frac{f_2^2}{V^{1/2}} \qquad [6\cdot18d]$$

Here H is the peak-to-peak variation of the transverse disturbing field, f_1 the focal length of the objective, f_2 that of the projector, M_1 the objective magnification, and M_2 the projector magnification. The total

length of the line into which an object point appears to be drawn out is given by a vector sum of these four terms, the individual displacements having different orientation depending on the rotation produced by the two lenses; d_2 and d_3 are of opposite sign. It is seen immediately that the second term, that is, the effect on the ray path between the objective and the intermediate image, is by far the largest, so that along this portion of the ray path shielding is most essential. Assume, so as to obtain a quantitative idea of the magnitude of the several terms, that $f_1 = f_2 = 0.4$ centimeter, $M_1 = M_2 = 100$, and $V = 50{,}000$ volts. Then

$$d_2 = 0.0106H \text{ cm}$$

The next largest terms, d_1 and d_4, are only one-hundredth as large as this.

It has been brought out in section 4·9 that normally the short-time fluctuations of the horizontal field component in a laboratory, provided that disturbing electrical machinery is either shielded or placed at a sufficient distance, are of the order of 10^{-4} gauss. Thus, if the region between the objective and the intermediate image is surrounded by a magnetic shield which reduces the field intensity within it by a factor of 100, the total deflection becomes approximately

$$d = 3{\cdot}10^{-8} \text{ cm (3 A.U.)}$$

in the most unfavorable case in which the three significant displacements have the same direction. Even with a variation of the horizontal field component by as much as $3{\cdot}10^{-4}$ gauss the total resulting deflection is, thus, less than 10 A.U.

It is obvious that increasing the shielding factor in region 2 materially above 100 is useless as long as regions 1 and 4 remain unshielded. If the shielding is accomplished by a single cylindrical sheath of radius r and thickness $d (d \ll r)$, the transverse field component inside the shell is weakened as compared to that in the absence of the shield by a factor (Eq. 14·66)

$$\eta = \frac{2r}{\mu d} \qquad [6{\cdot}19]$$

μ being the permeability of the material of the shield at low field strengths. For Permalloy (Fig. 14·11), for example, this permeability has a value 10,000, so that, for a shield radius of 2 centimeters, a reduction factor of 100 is obtained if the thickness of the shield is made only 0.4 millimeter.

In general, it is desirable to use a considerable factor of safety, to take

account of the weakening of the shielding action by the hole cut into the shield to permit the observation of the intermediate screen, as well as for chance variations in the permeability of individual shields. This may be accomplished, without an increase in the amount of material employed, by subdividing the shield into several coaxial cylinders spaced a certain distance apart. Thus if, in place of the single cylinder, two cylinders of radius 2 centimeters and 1.5 centimeters, each 0.2 millimeter thick, are employed, Eq. 14·68 shows that the shielding factor becomes $6.4 \cdot 10^{-4}$, which contains the desirable safety factor. If the initial permeability of the shielding material is only 3000, the shielding factor of the double cylinder becomes $6 \cdot 10^{-3}$, as compared to $3.3 \cdot 10^{-2}$ for the single sheath. It is seen that, with material of high initial permeability, a very light shield is sufficient to render ordinary external magnetic field fluctuations harmless.

If relatively large voltage fluctuations are present in the electron microscope, as may easily be the case in the electrostatic instrument, a stationary transverse field H_o, such as the horizontal component of the earth's field, has the same effect on the image as a fluctuating field of peak-to-peak amplitude $(\frac{1}{2})H_o \Delta V/V$, where ΔV is the total variation of the operating voltage. For example, if the voltage fluctuates by 1 per cent and if the horizontal field component is 0.2 gauss, the equivalent oscillating field has an amplitude of 10^{-3} gauss, ten times as great as the usual short-time fluctuations of the magnetic field. Thus, unless an electrostatic microscope is provided with a well-filtered voltage supply, it is necessary either to provide a superior degree of shielding along the entire instrument or to eliminate a large fraction of the transverse field component with the aid of carefully adjusted compensating coils.[8]

6·7. X-ray Protection. Wherever electrons strike matter, x-rays are generated. Since excessive exposure to x-rays is harmful to the human organism, giving rise to anemia, burns, and serious permanent injuries, precautions must be taken to reduce the quantity of radiation reaching the observer below a safe maximum. This is accomplished by interposing between the source of the x-radiation and the observer materials which will absorb most of the x-rays. For low operating voltages (less than 20,000 volts) the x-rays are so *soft* (readily absorbed) that the wall of the vacuum vessel as well as the air outside of it fulfil this function quite adequately. For the higher voltages used in the electron microscope added x-ray protection is needed.

The generally accepted measure of the x-ray dosage, that is, of the physiological action of a given amount of x-radiation, is the ionization produced by the x-rays in passing through a given quantity of air. The

[8] See Mahl, reference 3.

röntgen (r) is the x-ray dose which will liberate one electrostatic unit of charge ($3.3 \cdot 10^{-10}$ coulomb) in passing through 1 cubic centimeter of air at a temperature of 20° C and a pressure of 760 mm Hg. The maximum safe exposure to x-rays is taken to be 10^{-5} r per second for 200 working hours per month.[9]

The x-ray dose increases linearly with the electron current and with the atomic number of the target emitting the x-rays and somewhat more rapidly than linearly (in the range of 30 to 70 kilovolts) with the applied voltage. It decreases, furthermore (if the source is small in extent), as the inverse square of the distance of the observer from the source. The absorption of x-rays by matter, per unit mass thickness, however, increases approximately as the fourth power of the atomic number of the absorber and as the third power of the effective wave length of the x-rays, which in turn is inversely proportional to the applied voltage.

The x-ray emission from any target is highly inhomogeneous. Hence the dosage diminishes very rapidly initially as the thickness of the absorbing material is increased, the first layers absorbing practically completely the long wave-length components of the radiation, and more slowly afterwards. Figure 6·5[10] shows the thickness of lead required to reduce the dosage to which an observer is exposed at a distance of 1 foot (30 centimeters) from the tungsten target of an x-ray tube operated with 1 milliampere beam current and at voltages ranging from 30 to 75 kilovolts to 10^{-4}, 10^{-5}, and 10^{-6} r per second. It is seen that 1.1 millimeters of lead is adequate to reduce the radiation to a permissible value (10^{-5} r per second) at 70 kilovolts.

The thicknesses of various materials equivalent to this amount (1.1 millimeters) of lead — at 100 kilovolts — are shown below:[11]

Lead glass	6 – 9 mm
Lead rubber	2.5– 4 mm
Steel	7 mm
Barium sulfate plaster	8 –22 mm

For brass, such as is used to a large extent in the construction of the electron microscope, a slightly smaller thickness than that of steel (7 millimeters) affords equal protection.

In the electron microscope the total beam current is only half as large as that assumed in the above example, and, furthermore, the target materials are of lower atomic number than tungsten, so that a smaller amount of shielding would be sufficient. In particular, the shielding

[9] See Mutscheller, reference 4.
[10] See Bode and Glöde, reference 5.
[11] See Clark, reference 6, p. 148.

requirements become very small beyond the condenser and even less beyond the object, since here the original beam current has been reduced by a large factor. In the region between the intermediate image and the final screen it is usually only about 10^{-10} ampere. The only critical

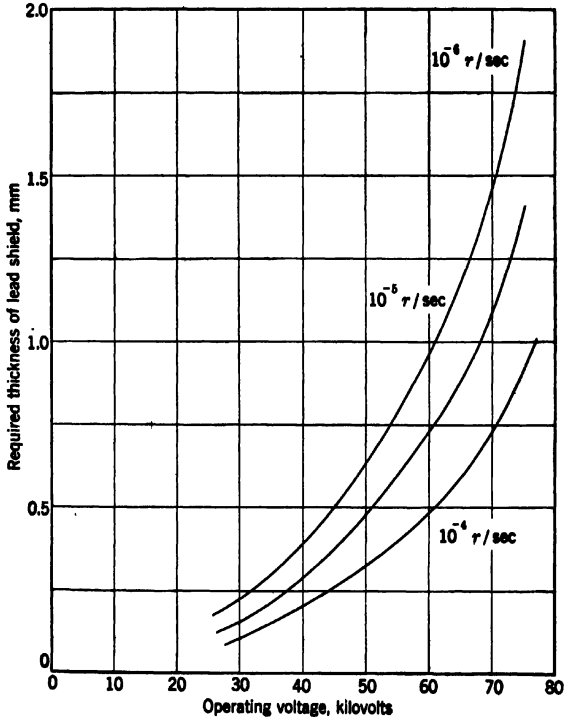


FIG. 6·5. Thickness of Lead Required to Reduce the X-ray Dosage at the Observer, 30 cm from the Tungsten Target of an X-ray Tube Operated with 1 ma Beam Current, to 10^{-4}, 10^{-5}, and 10^{-6} r/sec, respectively. (Bode and Glöde, reference 5.)

region is thus between the cathode and the condenser. Even here, up to 70 kilovolts, the wall of the microscope offers sufficient protection. It is only necessary to make the port window, for checking the beam alignment on an inclined viewing screen, of lead glass to be perfectly safe.

At higher voltages greater precautions must be taken. Since considerable penetrating secondary x-radiation is emitted wherever the primary x-rays strike, the entire gun region should be surrounded by

x-ray protection. However, here, also, the lower section of the microscope is adequately shielded by the microscope wall in view of the very small currents. To obtain an idea of the amount of shielding required near the gun, it may be brought out that, to reduce the dosage at an observer located a distance of a foot from a tungsten target bombarded with a 500-microampere beam to 10^{-5} r per second, it is necessary to place 4 millimeters of lead between the source and the observer at an operating voltage of 200 kilovolts and 15 millimeters of lead at an operating voltage of 300 kilovolts.[12]

With any electronic equipment employing very high voltages, such as those just mentioned, it is a wise precaution to explore the region surrounding the apparatus with an x-ray dosage meter.[13] A further precaution that may well be taken by the operator consists in carrying an x-ray dental film in his pocket. If, after the film has been exposed in this manner for two weeks, it shows definite fog on development, the shielding is inadequate.

REFERENCES

1. B. v. Borries and E. Ruska, "High-resolution microscopy with fast electrons," *Ergebn. der exakt. Naturwiss.*, Vol. 19, pp. 237–322, 1940.
2. J. Dosse, "Optical data of strong electron lenses," *Z. Physik*, Vol. 117, pp. 722–753, 1941.
3. H. Mahl, "The electrostatic electron supermicroscope," *Jahrb. der AEG-Forschung*, Vol. 7, pp. 43–56, 1940.
4. A. Mutscheller, "Physical standards of protection against roentgen-ray dangers," *Am. J. Roentgenol.*, Vol. 13, p. 65, 1925.
5. H. Bode and H. Glöde, "The production of x-rays in the operation of Braun tubes with high anode voltage," *Z. tech. Physik*, Vol. 20, pp. 117–124, 1939.
6. G. L. Clark, *Applied X-rays*, McGraw-Hill Book Company, New York, 1940.
7. W. Binks, "Nomogram for the determination of lead protection against high-voltage x-rays," *Brit. J. Radiol.*, Vol. 13, pp. 322–323, 1940.

[12] See Binks, reference 7.

[13] See Clark, reference 6, pp. 65–67.

CHAPTER 7

ELECTRON MICROSCOPE POWER SUPPLIES

7·1. General Requirements. All electron microscopes require a high-voltage source which possesses a certain degree of stability. Types utilizing electromagnet lenses demand an extremely stable source of high potential and, in addition, require one or more current supplies possessing the same or an even greater degree of stability. In Chapter 6 the maximum variation which may be permitted in any of these supplies has been calculated on the basis of a resulting image diffusion of not more than 10 A.U. The calculations gave the following percentages for the maximum permitted variations in the various supplies necessary for a conventional magnetic electron microscope: 0.0055 per cent for the objective lens (Eq. 6·15*b*), 0.013 per cent for the projection lens (Eq. 6·15*c*), 0.1 per cent for the condenser lens, and 0.011 per cent for the accelerating potential (Eq. 6·15*a*). The variations in the supplies need be less than the stated percentages only during the period between the final focusing of the instrument and the end of the exposure time. This interval is usually quite short, of the order of 30 seconds. It is, however, very desirable to obtain maximum long-time stability in order that the focus can be found readily and can be re-established with a minimum of adjustment. It is also desirable that the *regulation* of the high potential be good, that is, that changes in electron current, corona leakage, etc., result in the least possible changes in the voltage.

Electron microscopes using electrostatic lenses require much less stabilization of the high-voltage supply. In section 6·4 it is shown that the maximum variation which may be permitted in electrostatic instruments using 50,000-volt electrons is about 0.7 per cent. Such a degree of stability is easy to obtain in a conventional high-voltage supply such as that described in section 3·3.

7·2. High-Potential Supply. The accelerating potential required for an electron microscope varies according to the particular type of specimen being investigated and may have any value from a few to several hundred kilovolts. At present the range of 30 to 300 kilovolts has been explored, as indicated in Chapter 5. Most researches carried out with the larger, standard instruments utilize accelerating potentials between 50 and 100 kilovolts. The newer, compact electron microscopes em-

ploy lower voltages, generally in the range of 30 to 40 kilovolts. Voltages of several hundred kilovolts, finally, find application in special research with special instruments. The maximum beam current which the high-voltage source must supply is usually 1 milliampere or less.

The accelerating potential supplies for early electron microscopes were quite elementary. They consisted of conventional high-voltage step-up transformers and rectifiers of the type used in direct-current x-ray generators. The ripple of the output voltage was reduced by a

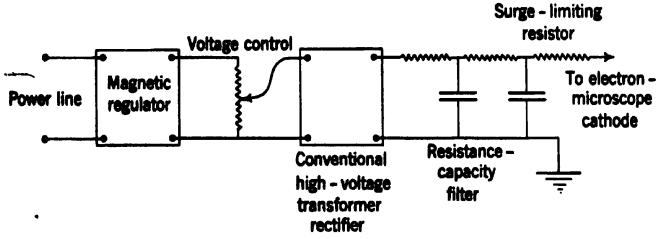


FIG. 7·1: Block Diagram of Conventional High-Voltage Supply.

resistance-capacity filter of one or more stages. In a supply of this type the output voltage depends directly on the primary voltage. Thus, to nullify the effect of line fluctuations on the sharpness of the image during exposure, the time constant of the filter would have to be made several times as large as the maximum exposure time. The consequent bulk and cost of the filters are such that this condition may well never have been fulfilled. Instead, at a later date, a line-voltage regulator of the saturated-core type was added. This reduced the line-voltage fluctuations by an order of magnitude and, hence, greatly increased the number of successful exposures. Even under these optimum conditions the elementary supply described remains bulky, cumbersome, expensive, and incapable of producing absolutely consistent results. Furthermore, the employment of the resistance-capacity filter results in poor regulation, that is, a close dependence of the operating voltage on the operating current. Also, the large amount of energy stored in the filter condensers may cause the flow of dangerously large currents on short circuit. Thus a voltage breakdown in the electron microscope, due to vacuum failure, may severely damage the instrument. Even so, most European electron microscopes, commercial models included, are powered by systems of this general type.[1] Figure 7·1 shows a block diagram of such a high-voltage supply.

A satisfactory solution to the stable-accelerating-potential problem

[1] See Ruska, reference 1.

is found in the electronically regulated rectifier using a high degree of inverse feedback. This method of regulation, well known in television and other electronic arts, operates by comparing a portion of the output voltage with some standard voltage such as that of a battery or glow tube, and by amplifying and applying the difference to control the current fed into the input. It renders the output voltage insensitive

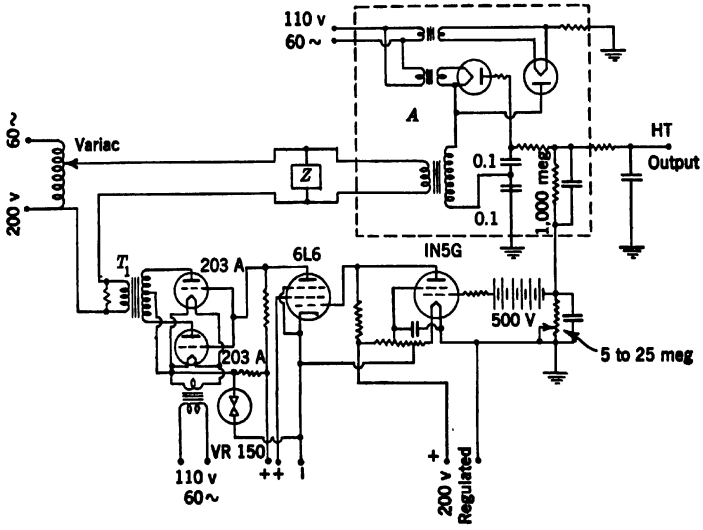


FIG. 7·2. Low-Frequency Excited High-Voltage Supply with Electronic Feed-Back Stabilization.

both to input voltage changes and output load changes. With large inverse feedback the output stability is limited only by the stability of the standard voltage and by drift conditions in the sampling and amplifying means. A 20- to 100-kilovolt regulated supply utilizing the inverse-feedback principle was developed in the RCA Laboratories in 1939.[2] The power was here supplied by a conventional x-ray generator for 20 to 100 kilovolts.

Figure 7·2 is a schematic diagram of this regulated high-voltage system, showing only the essential components. The high-voltage rectifier is enclosed by dotted lines and is marked *A*. A 1000-megohm, 100,000-volt carbon resistor, forming the high-voltage portion of the voltage divider and shunted by a compensating condenser, is mounted in the generator tank under oil. The high voltage leaves the generator

[2] See Vance, reference 2.

through a shockproof cable and enters another oil-filled tank containing the resistance-capacitance filter shown. The wire-wound resistor element of this filter serves as a current limiter for the system and prevents serious overloading of the generator in case of breakdowns and discharges in the microscope. The capacitance of the filter is made up of an oil-filled condenser plus the capacitance of 50 feet of shockproof cable leading from the high-voltage vault. The lower-voltage portion of the voltage divider is adjustable between 5 and 25 megohms for a voltage range of 100 down to 20 kilovolts and is used as the sample voltage for the regulator. This resistance appears at the remote-control position and is shunted by a condenser in order that the capacitance ratio of the divider approximate its resistance ratio.

As shown in Fig. 7·2, the sample voltage is opposed by the standard voltage, provided by a 500-volt dry battery, and the difference is impressed on the first tube (Type 1N5–G) of a direct-current amplifier system. This tube is direct-current-connected to a Type 6L6 direct-current amplifier actuated by a floating direct-current supply. The last tube drives, in parallel fashion, the grids of two Type 203–A power tubes, the plates of which are connected in push-pull to the secondary of a plate transformer. The primary of the plate transformer is shunted by an impedance-limiting resistor and is connected in series with the alternating-current input supply to the main generator. The Type 203–A tubes act as a variable resistance across the secondary and thus vary the primary impedance as the common grid voltage changes. Since the grids of the Type 203–A's are driven somewhat positive when the lowest impedance is attained, it is necessary that the grids be isolated from each other by the resistors shown in their grid circuits. These lessen the effect of the grid current in the tube whose plate voltage is, at the moment, negative on the positive swing of the grid in the other tube. The Type VR 150–30 glow tube serves to maintain the cathodes of the Type 203–A's at 150 volts positive with respect to the Type 6L6 cathode. The whole system is an inverse-feedback circuit with one point heavily by-passed. All the high-frequency attenuation is produced at that point. Since the loop gain of this system passes through unity at about 90 degrees phase shift and has no lower cut-off, it is inherently stable.

The network Z, connected across the input to A, serves to correct the power factor and keeps the input impedance of the generator within reasonable limits as the required output voltages are varied over a wide range.

The stability of the output voltage in the circuit shown in Fig. 7·2 is limited by a number of factors. On a straight loop-gain basis the

regulator should reduce the effect of line-voltage fluctuations several thousand times. Very fast line-voltage variations are reduced much less effectively, since the regulator is limited in speed by the power-line frequency. The standard battery has been measured to have less than five parts in one million variation within the 30-second period. However, in this unit the divider is oil-immersed and the circulation of the oil about the resistor causes the effective ratio of the voltage divider to vary because of the establishment of thermal and electrical gradients. The capacity divider, also, is affected by oil circulation. These last phenomena constitute the main causes of instability in this circuit. Drifting of the direct-current regulating amplifier is a further possible source of output variations. This effect has been reduced to negligible proportions by supplying plate, screen, and filament voltages for the input tube (Type 1N5–G) from a small electronically regulated power pack which maintains its voltage constant within less than 0.1 per cent at all times. Since the gain of this stage is about 100, the drifting in the following stages is negligible in effect. The peak of the ripple, or alternating component, of the output voltage, measured by means of a capacitance divider, is found to be 0.0016 per cent of the constant component. The regulator maintains the voltage across the voltage divider essentially independent of load so that the circuit impedance at the end of the resistance-capacitance filter is nearly equal to the value of the filter resistor. With proper control of corona and leakage at the microscope and storage-battery filament supply, no serious variations due to change in load current are observed.

Feedback-regulated supplies of this type, when operated from a fairly steady power line, may show variations of less than 0.006 per cent. This can be improved to the order of 0.004 per cent or better by wax-insulating the voltage divider. This system, therefore, provides sufficient stability for successful electron microscope operation. Its regulation, though superior to that of the elementary system, is, however, still limited by the resistance-capacity filter which lies outside the feedback loop. Furthermore, like the elementary system, it is bulky, expensive, and cumbersome. To a large extent, these and other drawbacks are inherent in the use of low-frequency (60-cycle) currents for generating the high voltages required. Thus all 60-cycle systems produce strong stray fields, from which, because of their low frequency, the microscope can be shielded only with difficulty. This circumstance makes impossible the combination of the microscope and its power supplies into a single, compact unit. The required large number of turns of the secondary and the impossibility of reducing wire size beyond a certain minimum render 60-cycle high-voltage transformers neces-

sarily bulky. Finally, the filter condensers must be large to reduce adequately the ripple voltage and other high-frequency fluctuations, which the regulator cannot compensate.

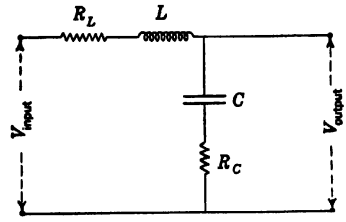


FIG. 7·3. Simple Resonant Circuit Corresponding to Quarter-Wave Line.

These several disadvantages may be removed by employing much higher frequencies, in particular radio frequencies, to drive the high-voltage rectifiers. Some of the gains resulting from this mode of operation, attained at the expense of a certain increase in the complexity of circuit design, are listed below:

1. Relatively small capacitors are required to reduce the ripple to a prescribed level. At the light loads involved, the principal component of ripple current flowing through the rectifier tubes is sinusoidal capacity current. The output shunt filter reactance can be made up advantageously of a network of coils and capacitors which presents zero reactance to the fundamental driving frequency and a low capacitative reactance to all the harmonics.
2. High impedance may be realized by low-loss resonant coils of minimum size and weight. This high impedance results in a minimum of exciting power and improves the overall power efficiency of the high-voltage supply.
3. Radio-frequency operation imposes negligible limitation on the regulator speed.
4. The radio-frequency fields of the radio-frequency circuits can be easily confined within the high-voltage generator, permitting the generator to be placed as close to the microscope as desired.
5. Only the simplest series-resonant circuit is required to produce the necessary voltage multiplication. Such a resonant circuit is shown in Fig. 7·3.

7·3. Radio-Frequency Circuits. The main criterion used for the determination of the operating frequency of a radio-frequency supply is the magnitude of the power losses in the coil due to eddy currents.

This power loss P is given by:

$$P = \frac{E^2}{Z} \qquad [7\cdot1]$$

where E is the voltage produced in the coil and Z is its resonant impedance, which is closely approximated by the expression:

$$Z = \frac{L}{C(R_L + R_C)} \qquad [7\cdot2]$$

Here L is the coil inductance, C the total capacitance across the coil, R_L the effective series resistance of the coil, and R_C the effective series resistance of the capacitance. C is maintained at a minimum and R_C is usually made small compared to R_L. The impedance Z is then essentially determined by L/R_L. At low frequencies R_L is very nearly equal to the direct-current resistance of the coil, and L/R_L is practically independent of frequency for a coil of given shape and size. As the frequency is raised a point is reached where R_L begins to increase owing to eddy currents and P increases. The operating frequency should be set at a point where there is only slight increase in coil loss due to eddy currents. Litz wire is usually employed to raise the operating frequency at which this slight increase in coil loss occurs.

In general, the higher the voltage required, the lower the operating frequency which is most suitable. The reason for this is that the capacity of high-voltage generators is necessarily larger, and, therefore, the conductor size in the coil must be larger. Since it is more difficult to reduce eddy currents in these larger conductors, a lower frequency is necessary to keep coil loss down. A list of high-voltage supplies of different developed alternating-current voltages, and their operating frequencies, follows:

	A-c voltage	*Rectifier circuit*	*D-c voltage*	*Frequency*
1	15 kv	Voltage doubler	30 kv	75 kc
2	17 kv	Voltage tripler	50 kv	75 kc
3	30 kv	Voltage doubler	60 kv	32 kc
4	50 kv	Voltage doubler	100 kv	24 kc
5	75 kv	Voltage quadrupler	300 kv	19 kc

The voltage ratio of the simple circuit shown in Fig. 7·3 is simply $2\pi fL/(R_C + R_L)$ or the Q of the circuit. This circuit may also be considered as a quarter-wave line with lumped constants. Such a line is called an impedance inverter because, for no dissipation, it has the characteristic that, if the input voltage is held constant, the output current remains constant, regardless of the load. It is of little consequence that in this circuit the output voltage is sensitive to changes in

the load, since the inverse feedback of the regulator reduces the overall sensitivity to load changes of the supply to negligible proportions. In fact, this impedance property is useful, in that it protects the other components of the system by causing the driving generator to unload and tend to approach an open circuit whenever the final output is short-circuited.

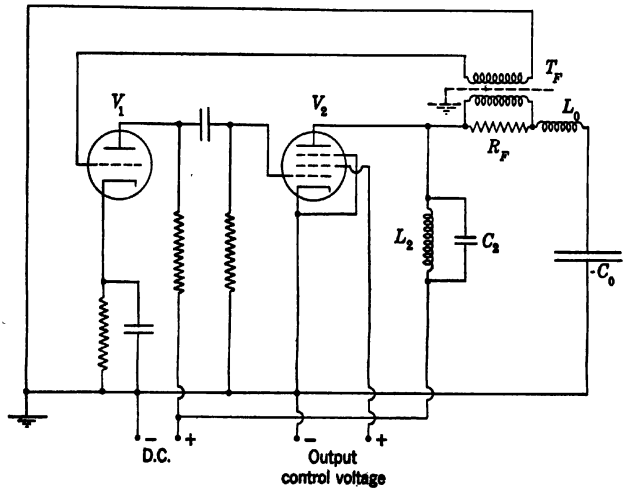


FIG. 7·4. Driving Oscillator for Series-Resonant Circuit.

The voltage ratio of this circuit varies critically with the driving frequency, which requires that the circuit be driven at its resonant frequency at all times. Since C consists of stray capacitance including that of the small rectifier, it usually has a high temperature coefficient, so that the resonant frequency changes considerably as the circuit warms up. This makes it desirable that the resonant load determine the frequency of the driving oscillator. Since the load to be driven is series-resonant — and not parallel-resonant, as usual — a special oscillator circuit must be used, in which the feedback to the input is proportional to the output current instead of to the output voltage. Figure 7·4 shows such an oscillator. It consists of a two-stage amplifier driving the series-resonant load L_0C_0 through the small feedback resistor R_F. The voltage across this resistor is fed to the amplifier input through the transformer T_F which has a broad frequency response and negligible phase shift over a broad frequency band centered on the resonant frequency of the load. The tank circuit L_2C_2 is parallel-resonant at

the approximate frequency of the load, but its L/C ratio is so high that it has negligible effect on the oscillator frequency. Hence, as the feedback gain is high enough and in the proper phase to produce oscillation only at or near the resonant frequency of the load, the circuit as a whole is constrained to oscillate at this frequency. The amplitude of the oscillation is controlled by the screen voltage of the output tube V_2.

Many variations of this circuit are possible and are used on occasion. For instance, the oscillation amplitude may be controlled by the plate voltage if desired and the oscillator circuit may contain one to three stages. The principle of current-controlled feedback, however, is utilized in every case.

7·4. Rectifier Circuits. As indicated above, a series-resonant input to the radio-frequency high-voltage supply is employed in anticipation of the use of a voltage-multiplying rectifier circuit instead of, for example, a simple half-wave rectifier. A number of considerations mark out multiplying circuits as most suitable for this application. Two of the most important among them are the questions of physical size and weight and of the power input required for a given output voltage and current.

If size is an important consideration and if the output voltage is quite high (60 kilovolts or more), the use of oil insulation is advantageous. This greatly reduces the difficulties of corona shielding and also permits smaller dimensions of various parts, such as the rectifier tubes. As compared with air-insulated units, oil-insulated units have the drawbacks of greater weight and greater fire hazard. In air-insulated units for moderate voltages the exaggerated size of air-insulated high-voltage rectifier tubes makes the employment of voltage-multiplying circuits and two or more small rectifying tubes an essential condition for compact construction. The fact that in voltage-multiplying circuits the maximum peak alternating-current voltage appearing anywhere in the circuit is limited permits further savings in space and leads to reduced stress on the insulation.

Voltage-multiplying circuits, as compared with a simple half-wave rectifier, also appear advantageous with regard to the power input required. Since, in the type of circuit considered, the whole rectifier system, regardless of type, is part of the series-resonant load, the exciting power[3] can be expressed in the form

$$\frac{E_o^2}{2Z} = \frac{E_o^2 CR}{2L} \qquad [7\cdot3]$$

[3] Since the output power adds the same amount to the input power required, irrespective of the character of the circuit, it is omitted in the above discussion.

where E_o is the peak input voltage, Z is the impedance of the resonant circuit, R is the series resistance, and C and L have the same significance as, for example, in Fig. 7·3. Since, for a given output voltage, E_o varies inversely as the number of multiplying stages and C increases linearly with the number of stages, while L and R remain substantially constant, it is obvious that the number of multiplying stages should be as large

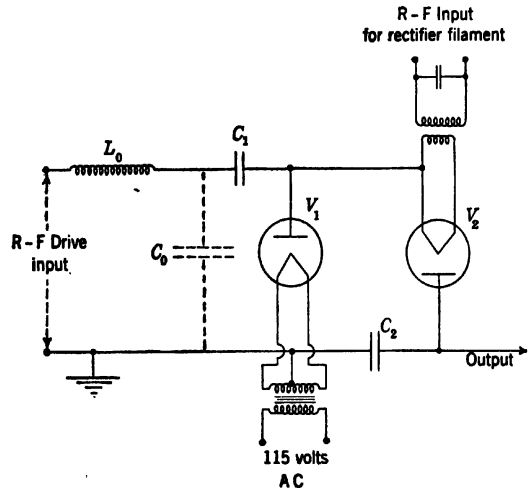


FIG. 7·5. Voltage Doubler Rectifier.

as possible. However, voltage-multiplying circuits require separately insulated filament power supplies, so that, in practice, the final decision regarding the number of multiplying stages must be made by balancing this extra complication against the saving in power, insulation, and space.

In the range of 50 to 100 kilovolts an oil-insulated doubling circuit is most commonly employed. Figure 7·5 shows a representative circuit for such a system. The first rectifier tube, V_1, in this circuit acts as a peak half-wave rectifier charging the capacitor C_1 to the peak voltage E_o of the radio-frequency input. The filament of the second rectifier tube, V_2, is connected to the capacitor C_1 and hence is at a potential consisting of the direct-current potential E_o and a superimposed alternating-current potential of peak value E_o. Thus V_2 also acts as a peak half-wave rectifier and charges the capacitor C_2 to the peak voltage which its filament assumes relative to ground, that is, $2E_o$.

It is of interest to note that the filament of the rectifier tube V_2 is also excited by radio frequency. The reason for this is that radio-frequency

transformers can be made with very low interwinding capacity, which further helps to conserve the exciting power required for the rectifying circuit (Eq. 7·3). Moreover, it is quite simple to attain the high-voltage insulation in radio-frequency transformers.

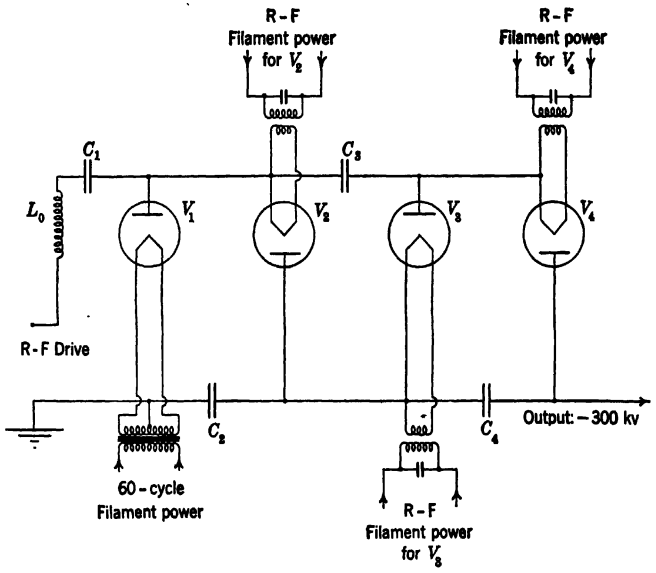


FIG. 7·6. Voltage Quadrupler for 300-Kilovolt Electron Microscope.

For lower operating voltages — for example, 30 kilovolts — a quite similar circuit, with air insulation in place of oil insulation, is suitable. For very high voltages, however, more complex multiplying circuits are desirable. Thus Fig. 7·6 shows an oil-insulated quadrupler circuit for the production of 300 kilovolts.[4] It is seen to consist of two voltage doublers in cascade with separate radio-frequency filament transformers and exciting oscillators for the rectifier tubes V_2, V_3, and V_4. Voltage-tripler circuits have recently found application in air-insulated 50-kilovolt supplies.

7·5. Regulating Circuits. The electron-microscope high-voltage supplies described above would, in themselves, be satisfactory for operation with electrostatic instruments. However, if they are to be used in conjunction with the magnetic type of electron microscope it is neces-

[4] See Zworykin, Hillier, and Vance, reference 3.

sary to add some regulating circuit which maintains the output voltage within the narrow limits defined in section 7·1.

The schematic diagram of a typical inverse-feedback regulated high-voltage supply[5] is shown in Fig. 7·7. The rectifier circuit is of the

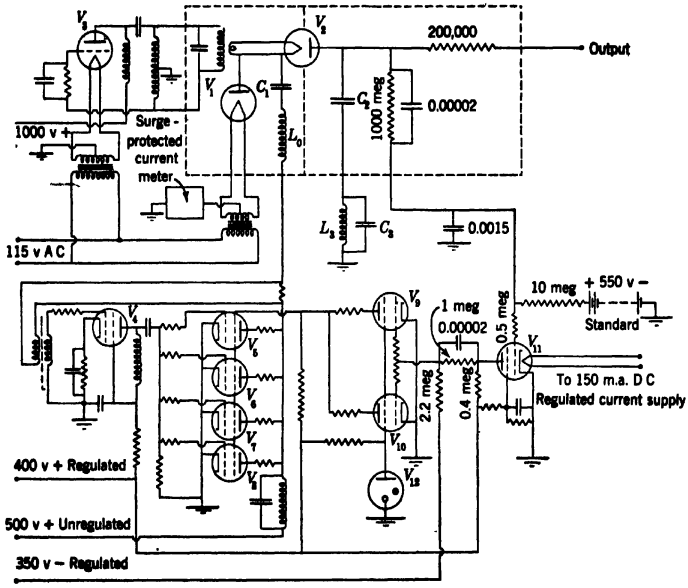


FIG. 7·7. Schematic Diagram of Radio-Frequency-Actuated, Feed-Back-Regulated, High-Voltage Supply.

doubler type, which has been described in the preceding section and which is shown in Fig. 7·5. It is driven by an oscillator such as that in Fig. 7·4. The sample voltage is obtained, as in the previous cases, from the low-voltage end of an output-voltage divider consisting of a 1000-megohm wax-enclosed carbon resistor and a wire-wound resistor of 10 megohms. This divider is shunted by capacitors so that the alternating-current and direct-current ratios are approximately equal. The difference between the sample voltage and the 550 volts of the standard battery is impressed on the grid of tube V_{11} which acts as a direct-current voltage amplifier. This voltage amplifier actuates a direct-current power amplifier consisting of the two tubes V_9 and V_{10} in parallel. The latter amplifier controls the screen-grid voltage of the

[5] See Hillier and Vance, reference 4.

driver-oscillator output stage (four tubes V_5, V_6, V_7, and V_8 in parallel), which, in turn, controls the power input to the rectifier circuit. The inverse-feedback loop formed in this way forces the sample voltage to be very nearly equal to that of the standard battery at all times.

There are a number of points of interest in the complete circuit, such as the following. The parallel arrangement of tubes in the power stages is a device which permits the use of a number of small, common radio tubes in place of single tubes which would have to be large and special in construction to meet the impedance and power-dissipation requirements. The resistance network connecting V_{11} to the following stage is arranged so that the current flowing from the negative 350-volt supply causes a voltage drop across the 1-megohm resistor sufficient to place the necessary negative bias on the tubes V_9 and V_{10}. The filament-heating power for the rectifier V_2 is supplied by the simple oscillator V_3. The 200,000-ohm resistor shown in series with the output acts as a current limiter in case of breakdown, protecting both C_2 and the microscope.

The components shown within the dotted rectangle are enclosed in a copper-lined steel tank and submerged in insulating oil. The alternating-current section of the tank is separated from the direct-current section by a copper shield which is also indicated by a dotted line. The second rectifier tube V_2 passes through a hole in this shield. In spite of this shielding there is sufficient capacity across V_2 to produce considerable ripple voltage. The filter circuit L_3C_3 is introduced in order to reduce this ripple voltage to a negligible value without the employment of an unreasonably large filter condenser C_2. The components L_3 and C_3 are chosen so that the effective inductance of the combination series-resonates C_2 at the driving frequency, providing a very low impedance path to ground for the ripple voltage. The circuit L_3C_3 is, in itself, parallel-resonant at a frequency between the fundamental and the second harmonic of the ripple frequency and thus acts as an inductance for the fundamental and as a capacitance for the second harmonic and all higher frequencies. Since the ripple is essentially sinusoidal, the harmonic content being low, the condition of series resonance is effective and reduces the ripple output to less than 0.0015 per cent of the output voltage.

The high-voltage supply shown in Fig. 7·7 and described above is capable of maintaining the output voltage steady to better than 0.005 per cent over the required time of 30 seconds. This circuit is, in essence, that which is used in the large RCA electron microscope (Model B). A photograph of the internal structure of the high-voltage oil tank of this instrument is shown in Fig. 7·8.

The small RCA electron microscope, which is described in section 5·1, possesses a high-voltage supply which furnishes 1 milliampere at 30 kilovolts with a maximum voltage fluctuation during an operating cycle of 0.005 per cent. The circuit of this supply is the same in principle as that shown in Fig. 7·7. Certain simplifications, such as the employment of air insulation and of a single-stage, single-tube oscillator, are made possible by the reduced demands on the circuit.

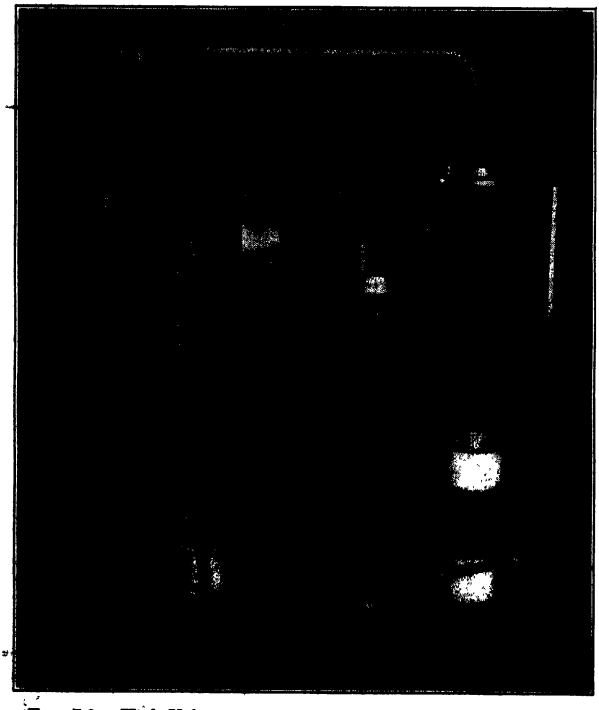


Fig. 7·8. High-Voltage Unit of the Type B Electron Microscope.

7·6. Magnetic-Lens Current Supplies. Early electron microscopes of the magnetic type used storage batteries to supply the exciting current for the lens fields. The current flowing through the lens windings was determined by the resistance of the winding itself and the external resistance which provided a means of control. In spite of the fact that storage batteries exhibit excellent voltage stability it was difficult to obtain satisfactory current stability because of the thermal changes of resistance of the system. The thermal changes arose from the heat

dissipation in the system and were difficult to eliminate because of the large bulk of the coil windings.

The required current stability can be attained quite simply by the use of a regulated current supply. Such a supply, employing conventional tubes, is shown in Fig. 7·9. In this circuit the voltage drop across the control resistance is compared with the voltage of a standard battery and the difference amplified and applied to the control grid of a power tube in series with the positive side of the main power supply, thus

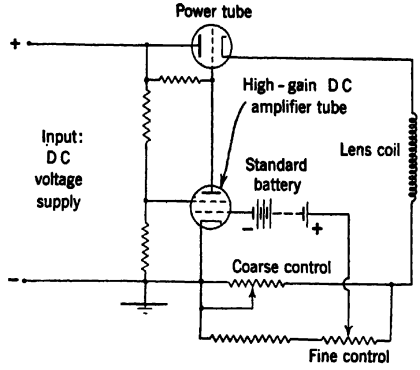


FIG. 7·9. Regulated Coil-Current Supply.

completing a powerful inverse-feedback loop. The main series resistor (eventually a number of resistors) is wound with wire having zero temperature coefficient. The rough control of the current is obtained in the case of the Type B electron microscope by varying this main series resistor in steps, whereas the fine control operates by slightly adjusting the fraction of the voltage drop compared with the standard battery. The output current of this type of regulator is almost entirely independent of the load resistance so that the whole system is insensitive to thermal changes caused by the current itself.

The direct-current input to the objective-lens current regulator is derived from conventional regulated low-voltage supplies. The heater current of the objective direct-current amplifier stage is also regulated so that it is steady to within 0.5 per cent. With these precautions the current in the lens coil can readily be held constant within 0.002 per cent. The single lens coil in the small RCA microscope is also excited by a circuit such as that shown in Fig. 7·9. Continuous controls are provided for the coarse and fine adjustment, which cover much smaller ranges than in the large instrument.

7·7. Filament Supplies. In early microscopes using hot cathodes to supply the electron beam batteries were used to heat the filament. The obvious inconvenience of insulating and controlling such batteries led to the development of alternating-current filament excitation. Exhaustive tests have shown that it is possible to use alternating-current excitation of an electron microscope filament with no detrimental effect on the image if certain precautions are taken to eliminate the effects of the alternating magnetic field which must be associated with it.

As an electron leaves the filament of an electron microscope it has a very low velocity and yet has to pass through the strongest part of the magnetic field associated with the heating current of the filament wire. As a result of these two conditions, the electron undergoes an appreciable deflection in a direction parallel to the axis of the filament wire. In an electron microscope using a direct-current excited filament this phenomenon appears as a constant displacement of the electron source, and compensation is made for it by realignment of the electron gun. However, in an electron microscope using an alternating-current excited filament the deflection of the electrons results in an apparent elongation of the electron source in a direction parallel to the axis of the filament wire. Optically this causes the effective aperture of the objective[6] to be elliptical with a resultant loss of resolving power in one direction in the image. If alternating-current excitation of an electron microscope filament is to be used, this effect must be reduced to the point where it has no observable effect on the image. This can be accomplished by the use of sufficiently fine filament wire (for example, tungsten wire with a diameter of 0.10 millimeter).

Figure 7·10 shows a microscope filament supply operating with 60-cycle current, such as is employed in the small RCA microscope. The filament current is supplied by an insulated step-down transformer from the 60-cycle line which is stabilized with the aid of a saturated-core voltage regulator. The transformer is provided with electrostatic shielding between the primary and secondary. Furthermore, the constancy of the potential of the filament tip, from which the electron beam is emitted, is assured by applying the high voltage to the center of a resistance shunting the filament. The diagram also shows the manner in which a bias is applied to the focusing aperture in front of the filament. The current from the high-voltage supply to the filament passes through a variable 100,000-ohm resistor. The voltage drop across this resistor is applied to the focusing aperture. Thus, for the maximum beam current of 1 milliampere the focusing aperture may be given a negative potential up to 100 volts.

[6] It is here assumed that, as is usually the case, the "effective aperture" of the objective is small compared to its physical aperture.

Even with the employment of a very thin filament the magnetic field of the filament circuit may have a slight effect on the beam in the region

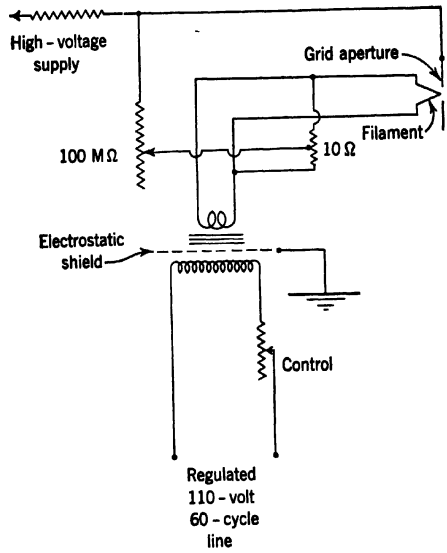


FIG. 7·10. Filament Circuit for Low-Frequency Excitation.

below the cathode aperture, resulting in an asymmetry of the effective source. This may be eliminated practically completely by the employ-

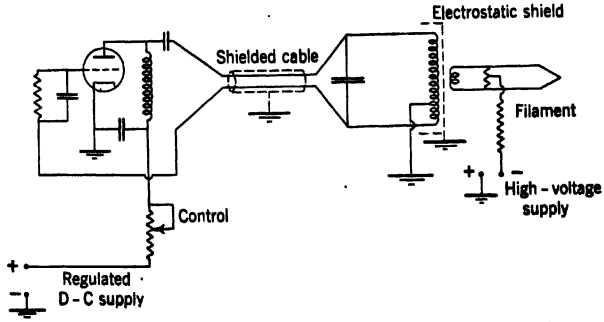


FIG. 7·11. Radio-Frequency Filament-Supply Circuit for Electron Microscope.

ment of radio-frequency excitation. Figure 7·11 shows a typical radio-frequency filament-supply circuit. Here the tank circuit of the driving oscillator is seen to constitute the primary of the step-down

transformer. With this arrangement considerable reactive power is stored in the transformer primary, enabling it to overcome the poor coupling with the secondary and to supply sufficient secondary voltage and current. At the same time it is not necessary for the leads of the transformer primary to carry the heavy circulating currents that would exist if the primary were only part of the tank circuit. Since a tuned secondary would require a large capacity, and hence would be expensive and cumbersome, an untuned secondary of several turns is employed. The heating current is controlled by varying the plate voltage of the oscillator and may thus, as for the low-frequency supply previously discussed, be conveniently adjusted by a resistor on the panel.

REFERENCES

1. E. Ruska, "Power and constancy of the current supplies for high-resolution electron microscopes," *Elektrotech. Z.*, Vol. 61, pp. 889–891, September 1940.
2. A. W. Vance, "Stable power supplies for electron microscopes," *RCA Review*, Vol. 5, pp. 293–300, January 1941.
3. V. K. Zworykin, J. Hillier, and A. W. Vance, "A preliminary report on the development of a 300-kilovolt magnetic electron microscope," *J. Applied Phys.*, Vol. 12, pp. 738–742, October 1941.
4. J. Hillier and A. W. Vance, "Recent developments in the electron microscope," *Proc. Inst. Radio Engrs.*, Vol. 29, pp. 167–176, April 1941.

CHAPTER 8

MANIPULATION OF THE ELECTRON MICROSCOPE

8·1. Preparation of the Specimens. Specimens must satisfy a number of conditions if they are to yield satisfactory images in the electron microscope. First of all, they must be thin enough to transmit a large fraction of the incident electrons with only minor changes in velocity and direction. For organic matter with a density of approximately unity the upper limit of the thickness at an operating voltage of 60 kilovolts is about one micron. This upper limit of thickness is inversely proportional to the density of the specimen. Greater specimen thicknesses may be tolerated if higher operating voltages are employed, as is brought out by the curves in Figs. 19·27 and 19·28.

Second, the specimens must maintain their character in vacuum and under the amount of electron bombardment essential for the formation of the image. Only the hardiest organisms (bacterial spores) have been observed in the electron microscope in a living condition.[1] Normally, specimens are freed of their volatile components before observation, signifying, in the case of bacteria and tissues, a removal of their water content. This does not, however, necessarily result in a basic change of the geometrical structure in the case of smaller entities.

The classes of objects that can be studied thus are: large organic molecules and viruses; finely divided matter, such as dusts, smokes, powders, and colloidal suspensions; substances forming thin films, such as soaps and plastics; bacteria; and, within limits, fine organic tissues and sections. In addition, the surfaces of etched metallographic specimens and other compact substances may be investigated indirectly by observing thin-film replicas of the surface.

Regardless of the character of the specimen, it must be mounted on a suitable support. Two types of specimen supports are in common use. The first consists of a disk, of some relatively inert metal, with a central aperture a few thousandths of an inch in diameter. This has the disadvantage that the specimen field which may be explored is rather limited. The other type, more widely employed at present, is free of this objection. It consists of a disk of fine-mesh metal screen, which may be either woven or electroplated. Most commonly the screen possesses

[1] See v. Ardenne and Friedrich-Freksa, reference 1.

approximately 200 meshes per inch. With this type of mount a large number of mesh apertures may be studied on a single specimen. In particular, if the specimen in one of the meshes is damaged by the electron bombardment, the observer can shift quickly to the next.

Some specimens require no support apart from this metallic base. Thus smoke particles caught on the wire-mesh disk by passing it over a flame tend to form chains reaching out from the edges of the mesh apertures toward the center of the holes. Other objects suitable for study form thin films covering the mesh openings. Plastics are examples of these. The metal disk need merely be dipped in a solution of the material having the proper concentration. After withdrawal, the mesh openings will be sealed by a liquid film, which solidifies as the solvent evaporates.

More frequently, the specimen requires a support in addition to the mesh wires. This support is most commonly provided by a strong, thin film, which, within the limits of resolution of the electron microscope, is structure-free. Numerous organic preparations are available which will form films with the required characteristics. Among them collodion and polyvinyl formal have gained the widest acceptance. Both of them yield structure-free films only about 100 A.U. in thickness, which have the requisite mechanical strength and are not readily destroyed by normal electron bombardment. A third supporting film, which is not quite so free of structure and has not quite so low a mass thickness as the collodion and polyvinyl formal films, is the aluminum oxide film. Its value resides primarily in its extraordinary stability under bombardment and when subjected to high temperatures. These advantages are shared by silica films prepared by evaporation on microscopically smooth surfaces of a plastic such as polystyrene.[2]

A collodion film[3] suitable as supporting membrane for electron microscope specimens may be prepared as follows. A single drop of a well-filtered 2 per cent solution of collodion in amyl acetate is placed on the clean surface of carefully distilled water. When the water surface is touched, the drop spreads over a fairly large area and the amyl acetate evaporates, leaving the desired collodion film floating on the surface of the water. A check on the quality of the film is available in the character of the interference colors which are seen as the amyl acetate evaporates. These interference colors should appear uniform either over the whole of the surface of the film or in fairly broad circles. In the last

[2] For details of their preparation, see the paragraph on the preparation of silica replicas, p. 252.

[3] Such collodion films were first introduced by Kirchner (reference 2) as specimen supports in electron diffraction work.

stages of evaporation the thickness of the film is much less than the minimum which will give interference colors. However, when the evaporation is completed the remaining film may be detected by a difference in the sheen of the water surface.



FIG. 8·1. Removing the Object Support from the Water Surface.

After a film has been obtained on the surface of the water, a number of screen disks are placed around the surface, as shown in Fig. 8·1. These can be picked up in a circular loop from below, as shown in the same photograph. In order to leave undisturbed all the film except the part being used, a needle is run quickly around the edge of the loop, thus cutting out that portion of the film. As it is lifted out of the water, the wire loop carries a screen disk supported in its opening on the collodion membrane and a certain amount of water which still adheres to the lower side of the membrane and the loop. This is inverted and passed over a small cylinder with a flat top, as shown in Fig. 8·2. A pair of forceps is now used to lift the cylinder through the loop so that the loop may be taken off the bottom. After this operation the screen disk rests on the top of the cylinder and is completely covered by the collodion membrane to which adhere, usually, a few drops of surplus water. The surplus water is then drawn off immediately by means of a micropipette.

As water does not readily wet collodion, the water drawn off by the above method comes off very completely and takes all foreign material which does not adhere to the collodion membrane with it. The specimen mount is then ready for the addition of the specimen (Fig. 8·3), unless, indeed, the specimen was suspended in the collodion solution and is thus already embedded in the film. Collodion films prepared in the manner described are found to be extremely clean and free from structure appreciably above the limit of resolution of the electron microscope.

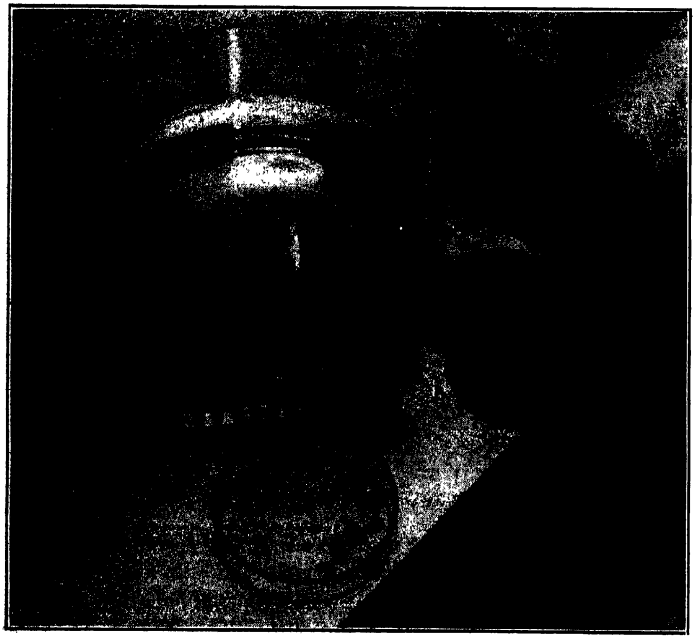


FIG. 8·2. Depositing the Object Support on the Cylinder and Removing Excess Water.

Films of polyvinyl formal,[4] which exceed collodion films in mechanical strength and durability, may be prepared by a procedure described by Schaefer and Harker.[5] A glass slide cleaned with magnesium oxide powder or chromic acid cleaning solution is dipped into a 0.1 per cent polyvinyl formal solution in dioxane or ethylene dichloride and with-

[4] Polyvinyl formal is marketed under the trade name Formvar 15–95 by Shawinigan Products Corporation, Empire State Building, New York City.

[5] See reference 3.

drawn, letting the solvent evaporate. After evaporation (consuming less than one minute) the film is scored with a needle around the edges and crosswise, so as to subdivide the film into areas of the required size. The plate is then held vertically over a tray containing water with a clean surface and quickly plunged under the water surface. After being submerged for about a minute the plate is withdrawn slowly, permitting the water to peel off from the surface, and then lowered again at a slower rate. The film will then float off on the surface of the water and can be utilized for forming an object support in the same manner as a collodion film.

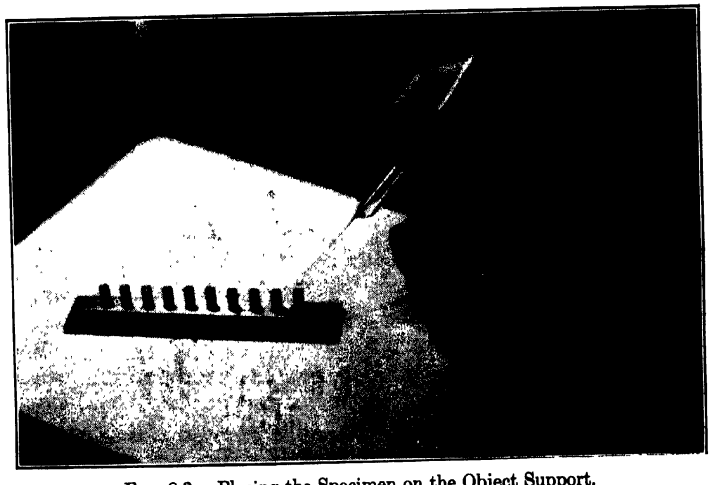


FIG. 8·3. Placing the Specimen on the Object Support.

Aluminum oxide films have been introduced as object supports by Hass and Kehler.[6] The following detailed instructions for preparing films 500 to 1500 A.U. in thickness are given by von Ardenne and Friedrich-Freksa.[7] A strip of aluminum foil, 1 to 2 centimeters wide and 5 centimeters long, is placed as anode in a 10 per cent sulfuric acid solution. The cathode is formed by a similarly shaped sheet of platinum placed at a distance of 2 centimeters from the anode foil. Employing a voltage of 10 volts, yielding an initial current of 0.01 ampere, an invisible aluminum oxide film, not yet showing interference colors, is formed in 1½ minutes. If thinner films are to be obtained, the forming voltage must be reduced. The aluminum strip with the film is then

[6] See reference 4.
[7] See reference 1.

immersed in saturated mercuric chloride solution, which, in a few minutes, dissolves the metallic aluminum. The mercuric chloride adhering to the film is removed by subsequent washing in alcoholic iodine-potassium iodide solution. Finally, the film is washed with alcohol and placed on the object support. The film appears transparent by transmission and slightly grayish or yellowish under vertical illumination. Since the film readily swims off the object support, it is often advisable to increase its adhesion by placing minute droplets of rubber cement on the edge of the object support. The electron diffraction patterns obtained from these films[6] indicate that they are amorphous.

Once the object support has been covered with one of these films, the specimen may be deposited on it either in a dry form or from a liquid suspension or solution. In the first case the material is simply dusted onto the film, collected on it by passing the object holder through an atmosphere[8] holding the substance in suspension, or evaporated or sputtered onto it in vacuum. If the material is suspended or dissolved in a liquid, a droplet of the liquid is placed on the holder, as shown in Fig. 8-3, and the liquid is permitted to evaporate. In the latter case great caution must be exercised to minimize the amount of foreign material present, since the dried residue of that material may greatly affect the image obtained. Also, care must be taken that the liquid is one which does not harm the supporting film.

The methods of carrying out the process of purification in the case of organic substances, such as bacteria and viruses, are highly specialized. Normally, the specimens are finally obtained suspended in distilled water which, if sufficiently pure, leaves no residue.

In the case of inorganic (colloidal) suspensions, both the choice of the medium and its hydrogen-ion concentration are of great importance, if the specimen is to be well dispersed. Several methods have been adopted for preventing the aggregation of particles on the film. Thus von Ardenne[9] places the specimens during the drying period on the prongs of a 100-cycle, electromagnetically excited tuning fork. The Columbian Carbon Company,[10] on the other hand, has found it advantageous to subject deposits of dry particles to a discharge in air of a Tesla coil (vacuum-leak tester) or to an air jet. Schaefer and Harker[11] deposit a monomolecular layer of a protein, such as insulin or egg albu-

[8] The addition of vapor of a solvent of the film material promotes adhesion of the particles to the film.

[9] See reference 5.

[10] See reference 6.

[11] See reference 3.

men, on a mounted film. This protein causes the aqueous suspension of the material to be studied to wet the surface evenly, preventing the formation of droplets, and consequent clumps of material, on the film during drying.

If the material is suspended or dissolved in a liquid which does not mix with water, such as collodion solution or oil, the suspension or solution may be spread on a water surface like a collodion supporting film and the spread film may be placed on a prepared mount in the same manner as the original supporting film. If the spread film solidifies and is mechanically strong, an additional supporting film becomes superfluous.

It is obvious that innumerable modifications of the techniques of specimen preparation here described may be employed to advantage in specific cases. A few of these are described briefly in Chapter 9, in connection with the electron microscope investigations in which they were used. Much additional information will be gained by reference to the original articles.

In almost all instances certain precautions must be observed to prevent rupture of the supporting film during bombardment by the electron beam. The consequent motion of the object during exposure is a frequent source of defective micrographs. Rupture results from two causes: the electrostatic forces between different portions of the specimen which charge up under bombardment and overheating of the supporting film caused by excessive energy absorption in the specimen. Both effects are most likely to be observed if the object material contains some very large particles or constitutes a very heavy deposit on the film. Finally, the employment of old collodion solutions in the preparation of the film may lead to rupture under bombardment as a consequence of reduced mechanical strength. A reduction of the beam current lessens the likelihood of film rupture in all cases.

A number of effective procedures are available, also, for the preparation of surface replicas of thick specimens, such as metallographic sections. The use of surface replicas was first suggested by Mahl,[12] who obtained, in particular, very fine pictures of etched aluminum specimens.[13] The polished and etched aluminum is oxidized anodically in a saturated borax solution with boric acid added. Then, after the surface has been subdivided into several sections with needle scratches, the specimen is immersed in saturated mercuric chloride solution. The mercury liberated at the scratches, where free aluminum is exposed to the solution, lifts up the sections of oxide film. These, after the rem-

[12] See reference 7.

[13] See Mahl, reference 8.

nants of metal adhering to them have been washed off with dilute hydrochloric acid, are ready for observation. The top contour of the oxide film follows the original etched surface, and the film is of uniform thickness (normal to the surface — Fig. 8·4) except insofar as the rate of formation of the oxide depends on the crystallographic orientation of the underlying surface. As a result the crystal surfaces appear darker in proportion as their slopes relative to the plane of the specimen surface are steeper. This gives the pictures a very plastic appearance.

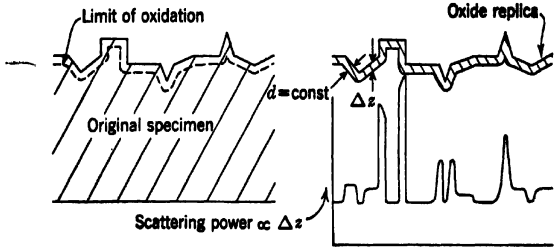


FIG. 8·4. Character of the Oxide-Film Replica.

Mahl has investigated a considerable number of other methods of forming surface replicas as well.[14] They include the formation of a thin plastic film on the surface to be studied, its oxidation by heating in the presence of air or oxygen, forming a thin oxide film, the deposition of a foreign oxide film by sputtering in an oxygen atmosphere, and the evaporation or electroplating of a thin metallic film on the surface. In all cases the film is separated by dissolving the base. Rüdiger[15] has successfully employed evaporated aluminum and beryllium films as replicas of etched steel surfaces. Beryllium films may be separated from the steel specimen by immersing the specimen for a few minutes in water slightly acidified with sulfuric acid. With the exception of the plastic film method, all these methods are applicable to a restricted range of specimens and result in replicas with "grain" or characteristic structure which may obscure the finer detail of the surface structure to be studied.

Among the more general methods of replica preparation three have proved particularly valuable. The first of these is the silver-collodion method.[16] The specimen, prepared in the same manner as for metallo-

[14] See reference 9.

[15] See Rüdiger, Bennek, Stäblein, and Volk, reference 10, and Rüdiger, reference 11.

[16] See Zworykin and Ramberg, reference 12.

graphic-light-microscope observation at high magnifications,[17] is placed in a vacuum chamber with one or several spirals of 15-mil tungsten wire loaded with small pieces of silver mounted above it. After a vacuum of the order of 10^{-4} mm Hg has been attained, current is passed through the tungsten filaments, whose temperature is raised gradually until most of the silver has evaporated. After the vacuum chamber has been permitted to cool, air is admitted and the specimen is removed. It should be coated with a film of silver thick enough to be readily pulled off the specimen with forceps after the edge has been scratched with a razor blade. The difference in the coefficient of expansion of the silver and the specimen often causes the surfaces to separate during cooling, facilitating this process.

Collodion solution is now flowed over the side of the silver film which was in contact with the specimen surface, care being taken — for example, by resting the bottom of the silver film on filter paper — that none reaches the opposite side. The concentration of the collodion solution should be such that the first interference colors (yellowish and then bluish) are barely visible. After drying, the silver film is immersed in 2–3 normal nitric acid, with the free silver side up. After a few hours most of the silver should be dissolved. Frequently, the film is found to float on the surface. It may now be washed by replacing the nitric acid by several changes of distilled water. Residues of foreign material may be removed by immersing the film in a solution of iodine in hydriodic acid.[18] The replicas so obtained are positive replicas. Ridges on the original specimen are reproduced as similar ridges on the collodion replica, valleys as valleys (Fig. 8·5). Surface tension causes the free surface of the collodion film to remain approximately flat over any small region of the specimen, so that the thickness of the replica becomes a measure of the elevation of the surface structure of the specimen over the average level of the surface. Accordingly also, with due allowance for the peculiar contrast conditions in the electron microscope,[19] the relative darkness of different parts of the image indicates the relative heights of the fine structure of the specimen surface.

A procedure which is incomparably faster than that just described, but requires considerably greater care and dexterity, has been outlined by Schaefer and Harker.[20] It constitutes, in a sense, a refinement of Mahl's plastic-film method. A negative polyvinyl formal

[17] This refers to the polishing and etching. The specimens should *not* be mounted in plastics, which may readily disintegrate during the evaporation process.

[18] Drs. Simard and Nelson of the Battelle Memorial Institute recommend 2 g of iodine in 5 cc concentrated hydriodic acid diluted to 1 part in 10 of water.

[19] See section 19·4.

[20] See reference 3.

replica is obtained by direct stripping from the specimen surface. Great care must be taken to have the surface completely clean. It may be impossible to remove the film from an etched metal surface if more than five minutes elapse between the drying after the etching of the specimen and the formation of the polyvinyl formal film. Accordingly, directly

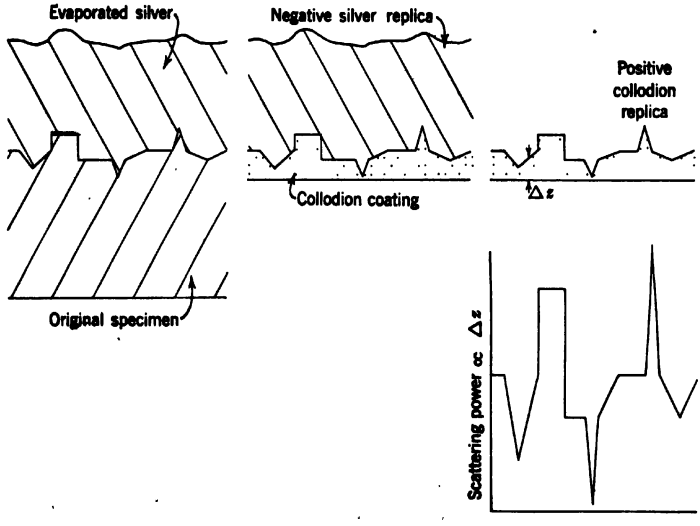


FIG. 8-5. Character of the Replica Obtained with the Silver-Collodion Process.

after the final polishing with magnesium oxide powder and the etching, the surface is dried quickly with a jet of filtered air and lowered face down into a petri dish containing a 0.5 per cent solution of polyvinyl formal in dioxane. The specimen is then swung up into a vertical position above the dish and permitted to dry. The resulting film should have a thickness range of 500 to 750 A.U., except at the very bottom, where the liquid accumulates. Then the specimen is immersed in water to a depth of about 3 centimeters and the thick end of the film is carefully teased until it piles up in a small roll. This rolled end is gripped with tweezers and the film is pulled toward the thin end. The separated film, or portion of the film, is swung up so that it reaches the surface in an extended condition, where it will shed water and float. The replica is now ready for mounting. This entire operation may consume only a few minutes.

Schaefer[21] has devised a number of methods of obtaining polyvinyl formal replicas which make unnecessary the immersion of the specimen in water and are, thus, particularly adapted for specimens which are corroded by water. In the first a 20 per cent solution of gelatin dissolved in warm water is applied to the surface coated with polyvinyl formal (about 0.5 drop per square centimeter). The gelatin then is dried with a stream of warm air. During the drying it either cracks off, taking the replica film with it, or may be split off with the aid of a razor blade. The film is placed, gelatin side down, in a tray and washed with several changes of warm water. After this treatment the replica film, floating on the surface, may be mounted.

Alternatively, a dilute, 0.5 per cent solution of gelatin is placed over the polyvinyl formal film on the specimen in a drop several millimeters thick. This is set on the freezing coils of a refrigerator. After freezing, the lower part of the specimen is placed in liquid air and left there until the gelatin and replica film split off because of unequal contraction of the materials. The film, including the drop, is lifted off and placed on filter paper. After being warmed to room temperature it is floated on a water surface and mounted.

In the last method, which is the simplest of all, moist air is breathed on the replica film coating the specimen and a disk of wire mesh is pressed immediately on it. Then a strip of cellulose Scotch tape is placed over the disk and the surrounding film and is pulled off. Under appropriate conditions the disk with the replica film over it comes off with the Scotch tape and, after being removed from the tape, may be examined in the electron microscope. In all cases, the polyvinyl formal replica being negative, ridges on the original surface correspond to valleys in the replica and, hence, bright portions of the electron microscope picture; depressions in the specimen appear dark.

A final process, described by Heidenreich and Peck[22] of the Dow Chemical Company, yields positive replicas of silica with excellent resolution and contrast. The specimen is placed in a molding press with enough polystyrene molding powder above it to give a final molding one-half inch thick. The mold is heated to 130° C before applying pressure. Then a pressure of 2000 to 5000 pounds per square inch is applied, and the temperature is raised to 160° C. The mold is now cooled at constant pressure until the temperature has dropped well below 80° C. The polystyrene is now sawed off the edges of the specimen. If the specimen does not readily jar loose from the remaining polystyrene molding, the excess of the specimen is sawed off and the remainder dissolved in an

[21] See reference 13.

[22] See reference 14.

appropriate acid. Any mineral acid (other than sulfuric acid) in a concentration less than 1:3, acetic acid, or a caustic may be used without injury to the polystyrene.

After careful washing the polystyrene replica is placed in a vacuum chamber with a pressure lower than 10^{-4} mm Hg. Silica is evaporated onto it from a conical tungsten filament made of 0.020-inch wire, into which quartz splinters are placed. The amount that must be evaporated depends on the roughness of the surfaces of the polystyrene block, on its distance from the filament, and on its total surface area. For a typical specimen of pearlitic steel and a separation between the apex of the filament and the surface of 6 centimeters, the proper amount to be evaporated is about 3 milligrams.

As the silica has a very high mobility on the polystyrene, it will cover all sides of the polystyrene block. Before the block is placed in the solvent (ethyl bromide), the film is removed from the bottom and the sides with fine emery cloth. Then the surface is cut up into sections about an eighth of an inch on the side and immersed in the solvent. The block is pushed against the bottom of the vessel to make it adhere there. Within five minutes the replica films of silica will float off. They may be caught on mounting disks, then dried on filter paper and washed in fresh ethyl bromide. After the final drying they are ready for observation in the electron microscope. Electron diffraction patterns indicate that the silica films are vitreous.

Not all these replica procedures are equally applicable to all types of specimens. Their sensitivity to heat, pressure, and vacuum, as well as the available equipment, may influence the choice. The results obtained with the various methods in the past may serve as a guide to one entering upon metallographic research with the electron microscope.

8·2. Alignment of Magnetic Electron Microscope. In the alignment of the electron microscope the operator adjusts the position of the various lenses until all their axes of symmetry coincide. The behavior of the intermediate and final images in response to variations of the various lens currents is used as the criterion of the adjustment.

For aligning the microscope, a convenient test object (for example, magnesium oxide smoke or carbon smoke adhering to the screen utilized as object support) is introduced into the specimen chamber. Adjustments are made until a sharp image is obtained on the viewing screen of the photographic chamber. If, now, the instrument is aligned perfectly, a change in the objective current will result merely in a rotation of the intermediate image — hence also of the corresponding final image — about its center, accompanied by a change in magnification and a loss in sharpness. A change in the condenser current is reflected in a

change in brightness of the image. With a misaligned instrument, however, a change in the objective current will cause the image to change in magnification and sharpness and to rotate about some point which may lie well outside the image field. As the objective current is reduced, the

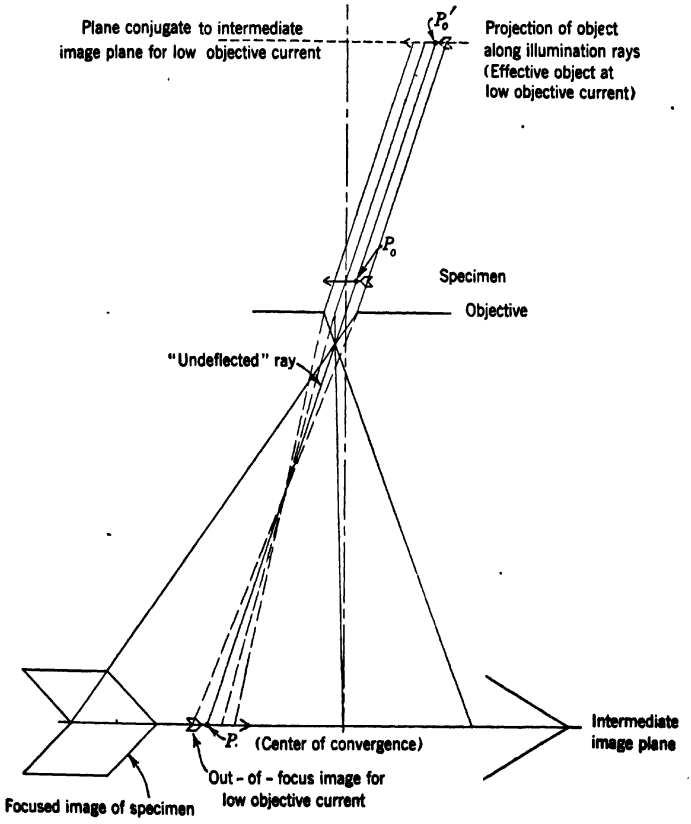


FIG. 8·6. Position of Center of Convergence with Reference to Out-of-Focus Image for Skew Illumination. (The Actual Rotation of the Images Is Not Represented on the Diagram.)

image contracts about the image point P_i (Fig. 8·6) corresponding to the object point P_o in which the illuminating ray aimed at the center of the objective intersects the specimen. The center of convergence P_i describes at the same time a circle whose center lies on the optic axis

of the objective, since the inclination to the axis of the ray passing through the center (or, more precisely, through the nodal points) of the lens is not altered by the lens.

Accordingly, the following is a suitable alignment procedure. Using very parallel illumination, that is, with the current in the condenser coil set well above the value at which the image of the filament is focused on the specimen plane, the final image is observed as the objective current is varied about the value for the best focus. In general, the image point will be seen to follow approximately circular paths about a center of rotation which lies outside the field of view. This center of rotation is not stationary, but shifts as the image is being focused. The first step consists in making the center of rotation for the sharply focused image coincide with the center of the field of view by a displacement of the objective adjustment screws. This adjustment is a first approximation to making the axes of the objective and projection lenses coincide. Unless the condenser lens is already accurately aligned, those points of the image which are at the center of rotation at exact focus will be seen to wander off the field of view as the objective current is varied. As can be seen in Fig. 8·6, the condenser can be aligned by displacing it in the direction which returns the out-of-focus image points to the positions which they occupied in the focused image. Since this adjustment will have changed the position of the center of rotation of the focused image, the first adjustment must be repeated. If the objective lens is accurately symmetrical and clean, alternating repetitions of the above two adjustments will quickly place the condenser, objective, and projector lens systems into accurate alignment. This condition exists when the point of the image at the center of rotation of the focused image does not move for any value of the objective current.

As these adjustments have been carried out at high condenser current and more or less arbitrary adjustment of the electron gun, it will be found on reducing the condenser current that the image of the electron source will not fall within the field of view of the objective. To correct this condition the following procedure is adopted. The intermediate screen is observed as the condenser current is reduced. Unless the electron gun is accurately aligned the illumination will be seen to become more intense but also to move off the field of view. It is then returned by the appropriate translation of the electron gun, which is repeated with further reduction of the condenser current until the image of the electron source is formed in the center of the field of view.

The focusing of the filament, while not truly a part of the alignment procedure, is nevertheless one of the most critical phases of the adjust-

ment of an electron microscope. This adjustment involves either a variation of the potential on the cathode shield or a mechanical adjustment of the position of the filament relative to the cathode shield in the direction of the axis of the instrument. In either case, the aim of this adjustment is to provide the smallest, most intense, and most symmetrical source for the electron microscope. The adjustment procedure is as follows: The negative potential on the shield or the position of the filament is adjusted until, with the appropriate values of the condenser and the objective currents, an enlarged focused image of the filament wire characterized by the die marks is observed on the intermediate screen. This adjustment is quite easily found and gives a reference point from which the final adjustment of the filament can be made. The filament is then moved farther away from the cathode shield or the shield potential is made more negative with respect to the cathode until the illuminated area of the specimen is about 50 microns in diameter, as determined from the intermediate image. This dimension is determined with the condenser current adjusted to make the illuminated area of the specimen a minimum. This method has been shown experimentally to give satisfactory illumination with maximum intensity and to be relatively insensitive to the electron gun configuration. The alignment of the gun is completed by adjusting the beam direction screws to obtain maximum symmetry of the specimen illumination and to eliminate any extraneous sources which may appear. The adjustments described above can best be accomplished with a large aperture in the objective and with the objective current reduced to about half its focal value.

8·3. Focusing. The focusing in the standard magnetic electron microscope is carried out by adjusting the objective current while observing the image of the specimen on the final-image screen. The accuracy of this focusing adjustment may be increased in several ways. Thus it is generally advantageous to augment the brightness of the image during visual observation by adjusting the condenser current so that the effective source is approximately imaged on the specimen. Apart from enhancing the brilliance of the image, this increases its sensitivity to small changes in the objective current, the depth of focus being reduced by the increased aperture of the imaging pencils. With the usual, relatively large, physical aperture of the objective the position of optimum focus is indicated by a slight reduction in contrast, contrast maxima being attained both at somewhat higher and somewhat lower objective currents (Fig. 19·18). This facilitates finding the exact focusing current.

After this current has been set, the condenser current is increased or

reduced so that the aperture of the illuminating pencils becomes substantially less than the effective aperture of the objective (normally about $3 \cdot 10^{-3}$ radian) and a picture may be taken. As indicated in section 4·3 the aperture of the illuminating pencils may also be reduced by mechanically exchanging the condenser aperture. Obtaining the micrograph at the reduced illumination and convergence of the illuminating pencils has two advantages: On the one hand, the employment of these narrower cones of illumination makes possible the attainment of the optimum resolution; on the other hand, it increases the depth of field and thus reduces effects of imperfect focusing. For an effective object aperture of $3 \cdot 10^{-3}$ radian the depth of field at the object which corresponds to a circle of confusion with a radius of only 10 A.U. is 0.7 micron. If the focal length of the objective is 3 millimeters the permissible variation ΔI in the objective current I leading to this degree of defocusing becomes, neglecting saturation effects, $\Delta I = 1.5 \cdot 10^{-4} I$. The saturation of the iron of the pole pieces increases this value considerably.

With objects which, themselves, have relatively low contrast, such as organic molecules or viruses on a thin film, the presence of heavy-metal or metallic-oxide particles greatly increases the ease of visual focusing.

If there is slight motion in the object, its effect on the image may be minimized by reducing the magnification after focusing[23] and, consequently, the required exposure time. This reduction in magnification may subsequently be compensated by light-optical enlargement of the electron micrograph.

For the attainment of the highest resolution in the image, it is advisable to take a series of exposures, changing the objective current in very small steps with the aid of a vernier adjustment. The best micrograph of the series is then selected for optical enlargement. With this procedure the sharpness of the images obtained becomes independent, within limits, of the accuracy with which visual focusing is possible.

An alternative practice consists of focusing the image visually under the same imaging conditions (illumination, focusing, and electron-optical magnification) as those prevailing during the exposure and employing light-optical magnification to bring out the finest image detail without loss in the apparent brightness of the image. This is applicable, in particular, to electron microscopes with fine-grain transmission screens which facilitate close examination with a high-power (for example, 20×) magnifier. The small RCA Electron Microscope described in section 5·1 is an example. The magnifier increases the aperture angle of the light

[23] This does not appreciably affect the sharpness of the image, as is shown in section 4·6.

pencils leaving the fluorescent image and ultimately passing through the pupil of the eye by a factor equal to the magnification and, hence, the amount of light reaching the retina from a particular section of the screen by the square of the magnification. Since this light is spread over an area larger than that covered by the same image portion without the use of the magnifier, the apparent brightness of the enlarged image is (except for unavoidable reflection and absorption losses in the magnifier) equal to the apparent brightness of the image on the fluorescent screen. With this procedure it becomes relatively immaterial whether the focus of the microscope is affected by changes in the illumination and magnification or not.

Essentially the same procedure, in a rather extreme form, is employed by von Ardenne[24] in his Universal Electron Microscope. Here a light microscope with a magnification of 100 is provided, which bears, in front of the objective, a minute right-angle prism coated on one side with a fine-grain fluorescent screen with a carefully ground single-crystal plate of activated zinc sulfide in the center. This fluorescent single-crystal plate, which is only about 0.001 inch thick, permits the resolution of separations in an image formed on it which are less than 0.01 millimeter. For focusing, this light microscope with its special fluorescent screen is pushed into the ray path a small distance below the projector, the vacuum seal being maintained by a metal bellows. Thus, although the gain in apparent brightness obtainable with a stage of light-optical magnification is fully utilized for focusing, the magnification of the micrograph is here much larger than that of the fluorescent image formed on the single-crystal screen.

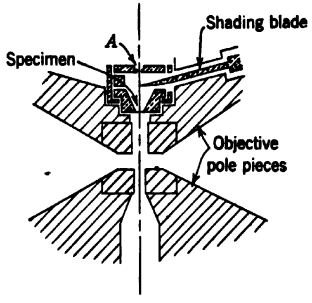


FIG. 8·7. Shading Arrangement for Protecting Sensitive Specimens during Focusing. (v. Ardenne, reference 16. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

If specimens which are very sensitive to electron bombardment, such as living substance, are to be micrographed, it may be desirable to carry out the focusing on an adjoining portion of the specimen, shading the remainder from the electron beam until the exposure is to be made. The arrangement employed by von Ardenne[25] for this purpose is shown in Fig. 8·7. The aperture *A* above the specimen restricts the illuminat-

[24] See reference 15.

[25] See reference 16.

ing beam to a small region of the specimen, protecting the latter from x-rays generated where the beam strikes the object holder, which, here, has a single opening. To begin with, the shading blade, which glides on an oiled plane, is moved up to the axis of the instrument, the object motion being adjusted so that the holder aperture is completely in the shadow of the blade. Then the specimen is translated carefully until a part of it is outside the shadow of the blade and the specimen is focused. Finally, the blade is withdrawn and the plate exposed, resulting in a picture containing an image of a shaded and an unshaded portion of the specimen. The mounting of the blade is entirely independent of the pole-piece specimen-holder unit, so that the motion of the blade does not communicate motion to the specimen.

8·4. Determination of the Magnification. A number of methods are available for the determination of the magnification of the electron microscope. The earlier ones measured, in general, the magnification of the two stages separately, the total magnification being then equal to the product of the magnifications of the individual stages. One of these[26] methods requires scales inscribed on the intermediate- and final-image screens and an object support consisting of a single centered aperture in a metal disk. The diameter of the latter is first measured in a light microscope. The ratio of the diameter of the luminous area on the intermediate screen to this measurement gives the magnification of the objective. Furthermore, the diameter of the aperture in the intermediate screen having also been measured by the light microscope, the ratio of the observed image diameter on the final screen to the above diameter yields the magnification of the second stage. A second method[27] determines, with the aid of a micrometer screw, the displacement of the specimen which corresponds to a given observed image shift on the intermediate screen. A second micrometer movement on the intermediate screen makes possible a similar measurement for the second stage of magnification. Although this second method is freed from certain limitations on the type of object support and the size of the intermediate image aperture inherent in the first method, both of them lead to total relative errors of the order of 20 per cent.

The method most widely used at present consists of taking micrographs of the same specimen with a light microscope and with the electron microscope at comparable magnifications. This is facilitated if a "short-mounted" object holder[28] is employed which increases the dis-

[26] See v. Borries and Ruska, reference 17.

[27] See Columbian Carbon Company, reference 6.

[28] Such as the specimen cartridge shown in Fig. 5·21 with the specimen in the upper, low-magnification position.

tance between the object and the center of the objective and thus decreases the objective magnification, increasing, simultaneously, the object field imaged. Thus electron micrographs with magnifications as low as 800 may be prepared with the Type B electron microscope. The ratio of the separations of two corresponding object points may then be determined on the two micrographs. The magnification of the electron microscope for the values of the objective and the projection coil currents and the operating voltage used becomes equal to this ratio multiplied by the known magnification of the light microscope.[29] The magnifications under all other operating conditions may then be determined by comparing corresponding object distances on different electron micrographs. The great sharpness of the electron micrographs makes such comparisons highly accurate. When the short-mounted holder is used, this method results in an error which may be as low as 1 per cent.

Another method capable of high accuracy[30] utilizes a replica, prepared by one of the procedures described in the first section, of a spectroscopic grating. So as to compensate for any conceivable scale distortion experienced in preparation of the replica, the grating constant is determined by measuring angles of diffraction obtained with the replica mounted in a spectrometer, both before and after taking an electron micrograph of it. The ratio of the line separations as measured on the electron micrograph to the grating constant is equal to the magnification of the electron microscope. Here also the accuracy of the method is increased by the use of the short-mounted holder.

8·5. Sensitive Materials. The photographic materials which are employed to record light images are also used to advantage for yielding a permanent record of electron images. The choice of the material, normally in the form of a plate or a film, depends primarily on two factors: its sensitivity and its resolution.

The sensitivity S of a photographic material to electrons may be defined as the reciprocal of the exposure E required to produce a prescribed density D, development being duly standardized. The density D is here the common logarithm of the ratio of the fraction of light transmitted by the exposed portion of the film or plate to the fraction transmitted by an unexposed portion. Exposure E, on the other hand, is defined as the product of the electron current density striking the sensitive surface and the time during which the surface is exposed. It is thus measured by the amount of charge that has reached unit area of the sensitive surface. Since, for high-speed electrons, the reciprocity

[29] The magnification of the light microscope is determined with an object micrometer.

[30] See Burton, Barnes, and Rochow, reference 18.

law is fulfilled within a very wide range,[31] the density obtained depends, for fixed electron velocity, solely on this product and not on the current density and the time individually. It is evident that for photographic materials having different shapes of curves representing the density as function of the logarithm of the exposure (H and D curves) the relative values of the sensitivity will depend on the choice of the standard density. For materials with similar H and D curves, this dependence does not exist.

The resolution of a plate or film may be defined as the least distance d_{ph} between two electron spots of infinitesimal extent for which their records on it, duly enlarged, still appear separated. This quantity depends on both the grain size of the photographic material and the spreading of the beam within the emulsion. The latter effect is a function of the velocity of the electrons and the thickness of the emulsion. For very slow electrons the penetration distance is so small that the spreading is, of necessity, slight. It increases with increasing velocity until the penetration distance becomes larger than the thickness of the emulsion. From this point on, the spreading may be expected to decrease again, since a progressively larger proportion of the electrons passes through the emulsion without large angular deflection. The efficiency of x-ray production is so low that x-rays produced in the plate have, normally, no appreciable effect either on the resolution or on the intensity of the image.

A suitable figure of merit F_m of a sensitive material employed to record electron microscope images is the ratio of the sensitivity S to the square of the least resolvable separation d_{ph}:

$$F_m = \frac{S}{d_{ph}^2} \qquad [8\cdot1]$$

This becomes evident from the following. Assume that the limit of resolution of the electron microscope for the type of object being examined is $d_{\min}$ and that the smallest separation which is readily distinguished with the naked eye is d_v ($\sim$0.2 millimeter). Then the useful magnification of the instrument is $d_v/d_{\min}$. A micrograph with this magnification is obtained by enlarging the plate, with the electron-optical magnification M_{el}, by the factor M_o:

$$M_{el} \cdot M_o = \frac{d_v}{d_{\min}}$$

Since, however, M_o must be no greater than d_v/d_{ph} if the limited

[31] See Bothe, reference 19.

resolution of the plate is not to have an adverse effect on the final picture,

$$M_{el} \geq \frac{d_{ph}}{d_{\min}} \qquad [8\cdot2]$$

At the same time, the exposure time T required to give a record of a given density is inversely proportional to the sensitivity S and to the current density at the image. Since the current density, in turn, is inversely proportional to the square of the electron-optical magnification,

$$\frac{1}{T_{\min}} \propto \frac{S}{(M_{el_{\min}})^2} = \text{const } S\left(\frac{d_{\min}}{d_{ph}}\right)^2 = \text{const}' F_m \qquad [8\cdot3]$$

The figure of merit is thus, quite properly, given by the reciprocal of the least exposure time which will yield an image, of prescribed density, containing all the information which the electron microscope can furnish.

Comparative measurements of the sensitivities to 60-kilovolt electrons of a large number of commercial emulsions, varying greatly in sensitivity, graininess, and contrast properties, indicate a rather small variation in the figure of merit. This is brought out by a visual comparison of pictures whose optical enlargement has been made inversely proportional to the square of the measured sensitivity of the original negative. Ordinary lantern slide plates were found to be among the most satisfactory of the materials tested. If high image contrasts are desired, commercial process plates may be found more suitable.

A considerable gain in the resolution, without an accompanying loss in the sensivitity, was found by von Ardenne[32] to be obtainable with plates on which the silver bromide grains are deposited without a binder (Schumann plates). The figure of merit of such plates is claimed to be thirty-six times as great as that of ordinary emulsions. The handling of Schumann plates is simplified if the silver bromide grains are deposited on gelatin instead of directly on glass.

If the sensitivity of plates for electrons at different voltages is compared, it may be expected that, for given current density, the rate of blackening will increase with the voltage until a large fraction of the incident electrons passes all the way through the emulsion. Beyond this point the laws of energy and current absorption (Eqs. 19·35a and 19·41) are such that a reduction in sensitivity with increasing electron velocity

[32] See reference 20. More recent measurements by v. Borries (references 21–23) indicate a much lower sensitivity for the Schumann plates. The discrepancy is to be ascribed in part to differences in the development of the plates (v. Ardenne, reference 24).

is to be expected. This has been confirmed by an experimental study of the sensitivity of Eastman Medium Lantern Slide Plates at operating potentials ranging from 40 to 212 kilovolts.[33] It was found that in this range the Hurter and Driffield curves had, within experimental error, identical shape. Thus the relative sensitivities at different voltages do not depend appreciably on the choice of the standard density. These curves are shown, with the exposure measured in coulombs (or ampere-seconds) per centimeter, in Fig. 8·8. Figure 8·9 indicates the variation of the sensitivity with the accelerating voltage.

In the development of the plates the same principles apply, in general, as in the development of light pictures. The curves of Figs. 8·7 and 8·8 were obtained with a development of 3 minutes at 19° C in Eastman D-72 developer diluted 1:2. A shorter development will, at the expense of a certain reduction in density, result in a plate capable of greater optical enlargement. Similarly, fine-grain development will improve the photographic resolution of the image, reducing the effective sensitivity of the material simultaneously.

At first sight it may appear surprising that the photographic plate, being bombarded by high-velocity electrons, does not charge up negatively. It is well known that at the voltages in question the secondary emission is inadequate to compensate the electrons absorbed in the material bombarded. However, when it is considered that the total current received by the plate rarely exceeds 10^{-10} ampere, it is seen that, even if the plate has a surface resistance of 10^{10} ohms, no part of it will charge up relative to the holder by an amount greater than a fraction of a volt.

8·6. Electron Stereomicroscopy. An electron micrograph represents a magnified projection of the specimen on a plane normal to the instrument axis, the object plane of the objective. Portions of the specimen lying outside this plane are still, within limits, imaged sharply, owing to the very great depth of field of the microscope. For the same reason, changing the position of the focal plane by quantities of the order of the specimen thickness does not result in recognizable differences in the sharpness of different portions of the object. In this respect the electron microscope differs from the high-magnification light microscope. The depth of field of the latter is only a fraction of its resolution, so that a photomicrograph images sharply only portions within and in the immediate neighborhood of the object plane. It is this fact that prevents obtaining satisfactory high-magnification stereomicrographs with the light microscope and necessitates the use of the method of optical sectioning for the determination of the three-dimensional structure of

[33] See Baker, Ramberg, and Hillier, reference 25.

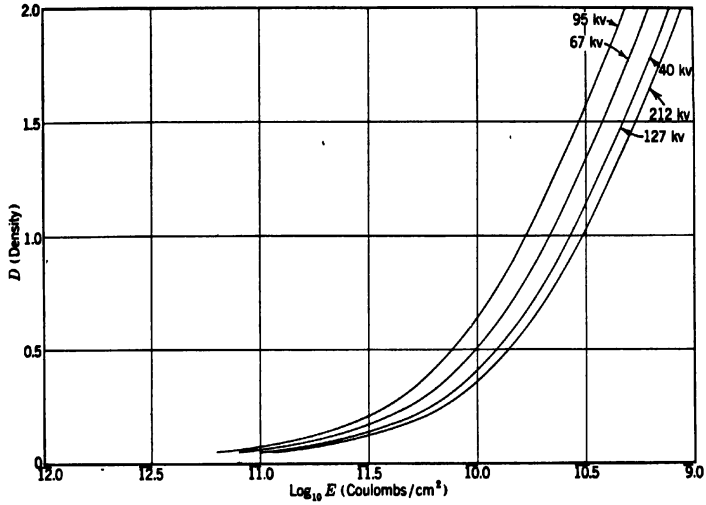


FIG. 8·8. *H* and *D* Curves for Eastman Lantern Slide Plates (Medium) at Different Accelerating Voltages.

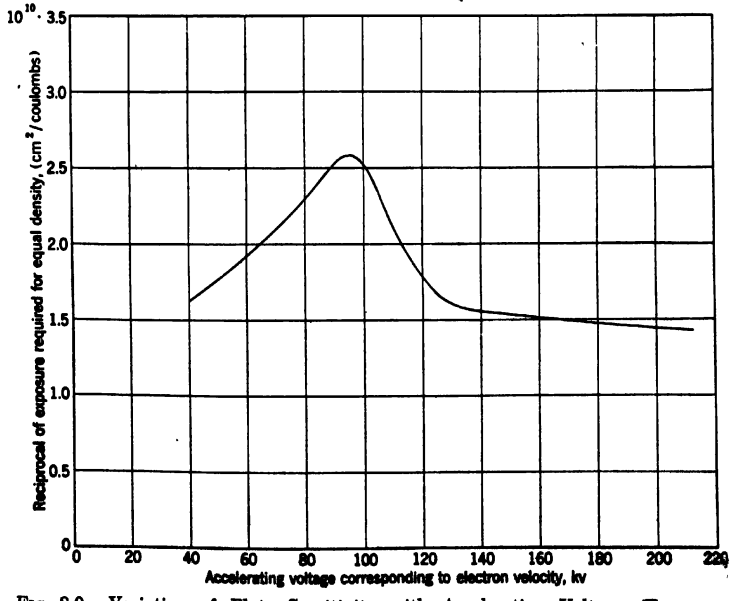


FIG. 8·9. Variation of Plate Sensitivity with Accelerating Voltage (Eastman Medium Lantern Slide Plates).

the object. In this method a series of micrographs are made with the objective focused in different planes of known separation. The three-dimensional structure of the object can then be reconstructed, though rather laboriously, from this series of micrographs.

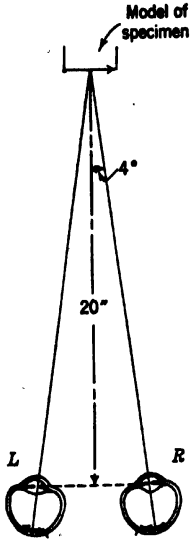


FIG. 8·10. The Convergence Angle in Binocular Vision.

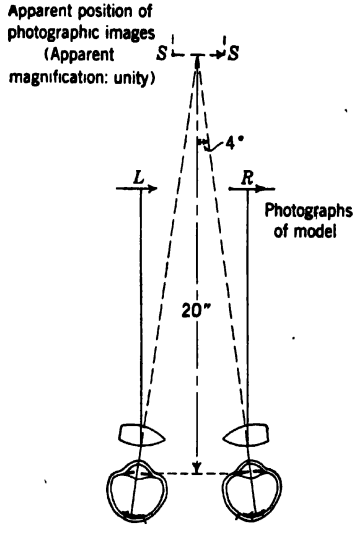


FIG. 8·11. Application of the Lens Stereoscope for Creating the Impression of a Three-Dimensional Image from Two Plane Images.

The great depth of field of the electron microscope, which makes this procedure ineffective, makes possible, in compensation, a utilization of the depth perception by binocular vision, which reveals the object in its true spatial relationships directly to the eyes of the observer.

In normal vision the three-dimensional appearance of objects results from the difference in the aspect which they present to the two eyes, or, in other words, from the effort of the brain to fuse two differing retinal images. A true idea of the three-dimensional character of an electron microscope specimen would be conveyed by viewing a model of the specimen, enlarged to scale by the total desired magnification and placed at some convenient viewing distance, say, 20 inches, from the eyes. Under these circumstances the angle between the two axes of the eyes — the *convergence angle* — would be 8 degrees (Fig. 8·10). Quite the same impression will result if, with the aid of a stereoscope, two images of

unity magnification, taken of the model by camera lenses placed at L and R, are presented to the two eyes, respectively (Fig. 8·11). Here the viewing lenses are provided with prisms which cause the eyes to view the two images superposed in the plane SS. It is evident that essentially the same effect will be obtained if two micrographs are prepared

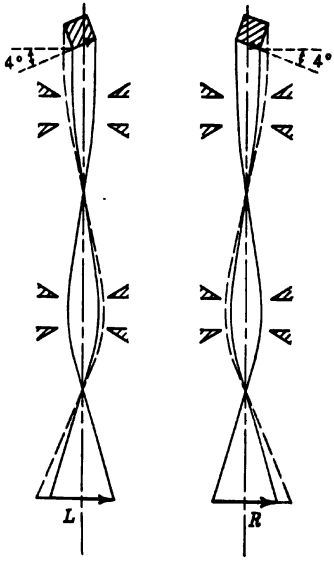


FIG. 8·12. Preparation of a Pair of Stereo Images by Tilting the Specimen Plane between Exposures.

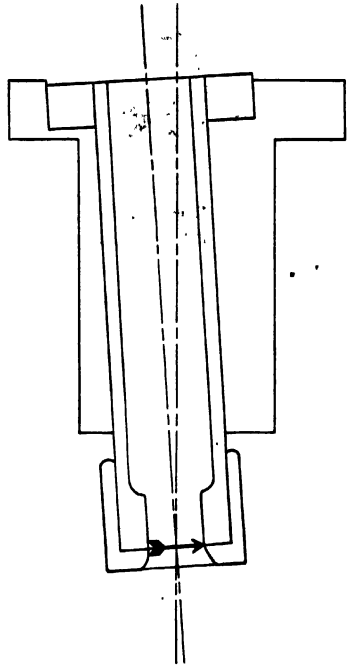


FIG. 8·13. Section through Specimen Holder for Preparing a Pair of Stereo Images.

of the original specimen as shown in Fig. 8·12, the normal to the specimen plane being inclined by 4 degrees in two opposite directions with respect to the instrument axis, and if these two micrographs are viewed in the stereoscope in normal fashion.[34]

[34] It should be noted that, since the illuminating beam effects a parallel projection of the specimen on a plane normal to the axis of the objective, the electron stereomicrographs are devoid of any perspective. In this respect the three-dimensional impression conveyed by them in the stereoscope differs from that obtained by viewing the magnified model directly. This lack of perspective does not in any way impair the perception of depth and simplifies the quantitative determination of the relative position of different portions of the specimen.

An object holder which facilitates the successive recording of the two required pictures is shown, in schematic section, in Fig. 8·13. The object is mounted at the bottom of the inner cylinder, whose axis is inclined by 4 degrees with respect to the axis of symmetry of the outer

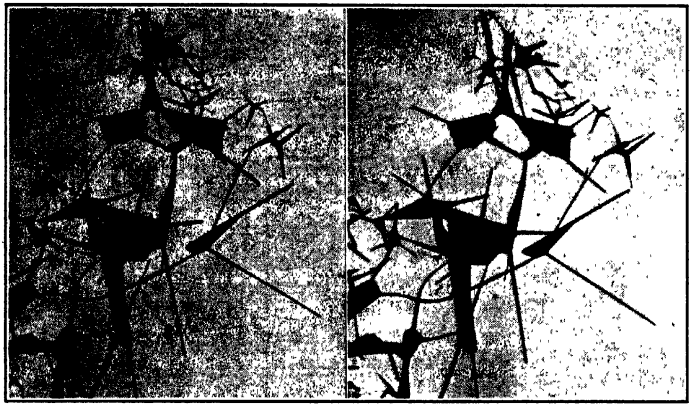


FIG. 8·14. Stereo Picture of Zinc Smoke (ZnO). Magnification 7600.

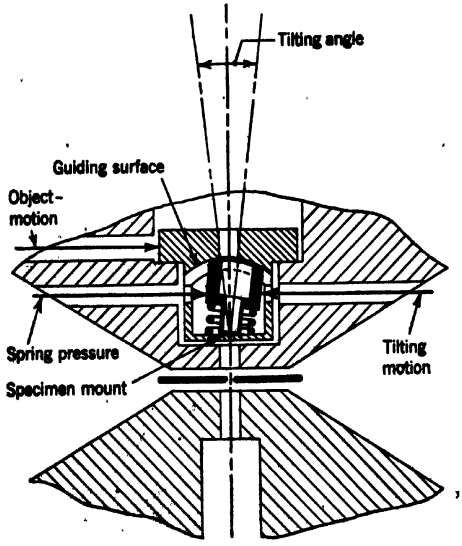


FIG. 8·15. Tilting Device of von Ardenne's Universal Electron Microscope. (v. Ardenne, reference 15.)

cylinder, the intersection of the two axes being at the center of the object. After one picture is taken the object holder is removed, the inner, inclined cylinder rotated through 180 degrees, and the object holder, also rotated through 180 degrees, is inserted into the object chamber. The specimen is now in its original position, except for a reversal of its direction of inclination with respect to the instrument axis. The sensation of depth conveyed by the resulting pictures viewed through a stereoscope is very striking. A typical pair of micrographs is shown in Fig. 8·14.

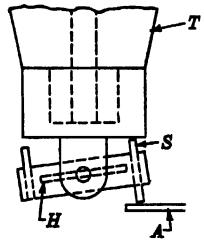


FIG. 8·16. Tilting Object Stage of Mahl's Electrostatic Electron Microscope. (Henneberg, reference 26.)

Several efforts have been made to carry out the tilting operation in the microscope, without removing the specimen holder. Thus von Ardenne[35] lets the upper end of the object holder, ground to a cylindrical or spherical surface whose center of curvature coincides with the center of the specimen, glide along a similar surface, a pusher serving to change the inclination with respect to the instrument axis (Fig. 8·15). A slightly simpler arrangement is employed by Mahl[36] in his electrostatic microscope (Fig. 8·16). The object support H is hinged to the object holder T. As the holder is moved down toward the objective beyond the normal focusing range, one end of the shaft S strikes the stop A and causes the support H to be tilted until the upper end of S strikes T. The angle of tilt can be adjusted in advance by displacing the shaft S. A third system, employed by Marton,[37] does not even require the tilting of the object (Fig. 8·17). The direction of the illuminating pencils is simply changed between exposures by altering the position of the condenser and, eventually, the electron gun. However, since this method makes use of off-axis portions of the objective, adequately sharp images

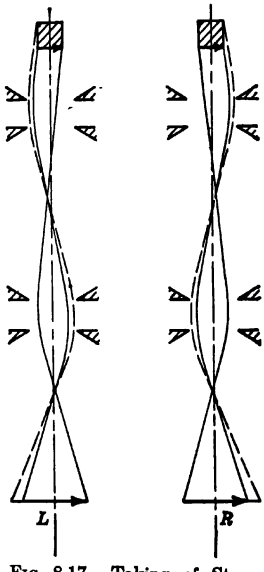


FIG. 8·17. Taking of Stereo Pictures by Changing Angle of Illumination. (Marton and Schiff.)

[35] See reference 15.
[36] See Henneberg, reference 26.
[37] See Marton and Schiff, reference 27.

can be obtained only for very small convergence angles, which yield an inadequate depth representation.

The taking of two stereoscopic pictures of the same object, the plane of the latter being inclined with respect to the plane normal to the axis by known angles $\theta/2$, also permits an accurate quantitative measurement of the relative elevations of different points of the object. Measurements of this kind have been made by Eitel and Gotthardt[38] on caolin

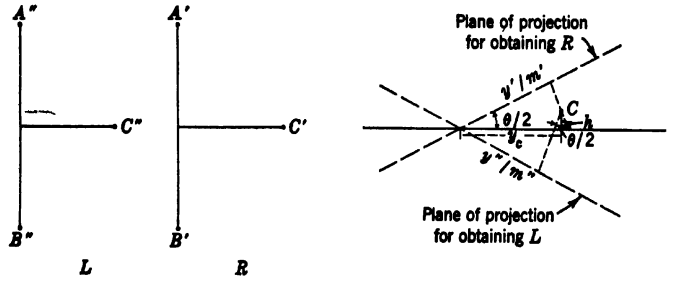


FIG. 8·18. Determination of the Coordinates of an Object Point from a Stereo Pair of Images.

crystals. Assume that L and R, Fig. 8·18, represent the pair of pictures obtained and that $A'B'$, $A''B''$ are the images of a line AB in the specimen which is parallel to the axis of tilt. Let m' and m'' be the image magnifications along these axes. Furthermore, let C' and C'' be the images of the object point C, which lies at a height h above the specimen plane containing the line AB. y' and y'' designate the distances of C' and C'' from $A'B'$ and $A''B''$, respectively. Then the geometric construction on the right, showing a section through C normal to the axis of tilt, indicates that the height h is given by

$$h = \frac{\dfrac{y'}{m'} - \dfrac{y''}{m''}}{2 \sin \dfrac{\theta}{2}} \qquad [8\cdot4]$$

The abscissa of the point C in the specimen plane is

$$y_c = \frac{\dfrac{y'}{m'} + \dfrac{y''}{m''}}{2 \cos \dfrac{\theta}{2}} \qquad [8\cdot5]$$

[38] See reference 28.

Equation 8·4 indicates that for precise measurements (large differences in the values of y'/m' and y''/m'' for given height h) a large tilt angle should be chosen, even though this will, eventually, limit the size of the sharply imaged object field in a direction normal to the tilt axis.

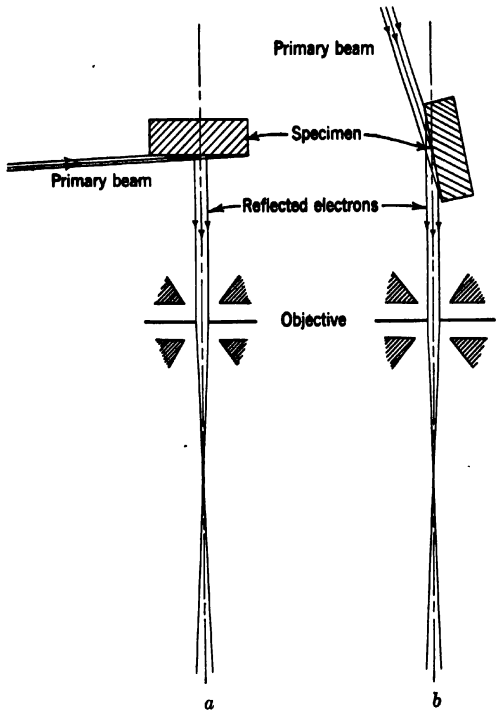


FIG. 8·19. Surface Examination with Reflected Electrons: (*a*) Specimen Normal to Microscope Axis (Ruska and Müller); (*b*) Specimen Almost Parallel to Microscope Axis (v. Borries.)

8·7. Surface Study with Reflected Electrons. A relatively minor modification of the standard electron microscope adapts it for the direct observation of the surfaces of compact objects with the aid of reflected electrons. Such electrons exhibit, normally, a great inhomogeneity of velocity, since most of them undergo considerable energy losses before they are scattered back out of the surface. These energy losses are minimized if attention is restricted to electrons which are incident on the surface at a glancing angle and leave it, similarly, at a glancing angle.

Figure 8·19 shows two arrangements which have been employed to

form images of surfaces by means of reflected electrons. The first[39] lets the electron beam be incident at right angles to the microscope axis on a specimen whose surface is almost perpendicular to the axis. The electrons are, thus, incident at a glancing angle, but leave the specimen in directions almost normal to its surface. The images so obtained are very faint and have resolutions inferior to those of high-magnification light microscopes.

In the second method[40] (Fig. 8·19*b*) the electrons are both incident on, and leave the specimen at, a very small angle — normally about 4 degrees — so that the inhomogeneity of the reflected electrons is reduced materially. It has been possible in this manner to achieve electron micrographs with resolutions along the direction of the axis of tilt which are considerably better than those obtained with a light microscope. The pictures, however, are very badly distorted; the magnification normal to the axis of tilt is only about one-fourteenth that along the axis of tilt, and the resolving power in this direction is reduced correspondingly. Practical importance can be ascribed to neither method, particularly in view of the superior results obtained with the replica techniques.

8·8. Electron Diffraction Patterns. The electron microscope picture gives the observer information regarding structural detail composed of large numbers of atoms. For knowledge of the more intimate molecular or crystalline composition of the individual particles he must rely on indirect physical methods. One of the most effective of these methods is electron diffraction. As discussed in section 19·9, any crystalline material will scatter electrons through discrete angles depending only on the lattice spacings of the material and the velocity of the electrons. In particular, if the material consists of randomly oriented crystallites, the diffracted rays will fall on cones of aperture angle θ, such that

$$n\lambda = 2d_s \sin\frac{\theta}{2} \qquad [8\cdot6]$$

Here λ is the wave length of the electron beam and d_s is the distance between two adjoining members of a particular family of lattice planes; n is simply an integer. Thus, if a plate is placed at some distance L below the specimen, normal to the axis of the incident beam, the ray cones will intersect it in circles whose radii R are related to the lattice constants d_s by

$$d_s = \frac{n\lambda L}{R}\left[1 + \frac{3}{8}\left(\frac{R}{L}\right)^2 - \frac{13}{128}\left(\frac{R}{L}\right)^4 + \frac{55}{1024}\left(\frac{R}{L}\right)^6 - \cdots\right] \qquad [8\cdot7]$$

[39] See Ruska and Müller, reference 29.

[40] See v. Borries, reference 30.

R/L being equal to tan θ (Fig. 8·20). Accurate values for the d_s corresponding to the several rings observed on the diffraction pattern permit the determination of both the shape and the size of the unit crystal cell of the material. Since the crystal structure is, in turn, a unique property of a given compound, electron diffraction effectively performs a microanalysis of the specimen. This is greatly aided by published lists of the lattice spacings d_s for a large number of common crystalline compounds.[41]

If, in place of a single electron ray, a parallel beam or a pencil diverging from a distant point source is available, a similarly sharp pattern may be obtained if an electron lens, placed below the specimen, images the source on the plate (Fig. 8·21). The constant coefficients in Eq. 8·7 will, however, be modified slightly and the rings, for equal distance between the specimen and the plate, be somewhat smaller in size. Under such circumstances it is most convenient to calibrate the diffraction camera by obtaining, with the electron velocity (accelerating voltage) to be used, a diffraction pattern for some substance of known lattice spacings and, hence, to plot the ring diameter against the corresponding lattice spacing (Fig. 8·22).

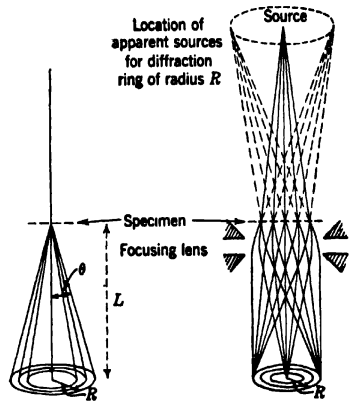


FIG. 8·20. FIG. 8·21.

FIG. 8·20. Formation of a Powder Diffraction Pattern with a Narrow Electron Beam.

FIG. 8·21. Formation of a Powder Diffraction Pattern with a Wide Electron Beam.

The sharpness of the diffraction pattern rings depends, apart from the regularity of structure of the specimen, on two instrumental factors: the uniformity of the electron velocity, which determines the value of the wave length λ, and the smallness of the divergence of the electron pencil illuminating any given point of the specimen. The divergence angle is given directly by the ratio of the diameter of the source to the distance between the source and the specimen. Since an extraordinarily homogeneous electron beam diverging from a very small source is available in the electron microscope, the microscope already contains the prime prerequisites of a precision electron diffraction camera. It is not surprising, in view of this, that a number of adaptations of the electron microscope for diffraction work have been carried out.

[41] See Hanawalt, Rinn, and Frevel, reference 31.

Possibly the most obvious arrangement leaves the specimen in its usual place and utilizes as point source either the crossover itself or a reduced image thereof formed by the condenser lens. The objective is employed to focus the pattern on the plate. If the pattern is to be recorded in the regular photographic chamber of the microscope it is necessary to remove the projector lens or to reduce the obstruction caused by it. In the Siemens microscope this is accomplished by

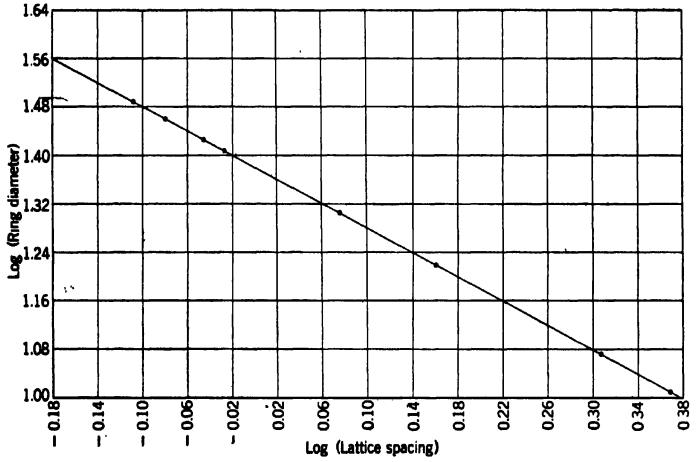


FIG. 8·22. Calibration of Type B Electron Microscope Diffraction Camera at 78 Kilovolts. (Courtesy *J. Applied Phys.*, reference 34.)

removing the projector pole pieces before recording a diffraction pattern (Fig. 8·23).[42] The microscope, with the pole pieces removed, is still capable of forming a survey picture of the specimen with a magnification of 1600. The principal drawback of this arrangement is that the projector pole-piece holder restricts the deflection angles θ which can be recorded to values less than 3.5 degrees.[43] So as to remove this drawback, as well as that of having to open up the microscope to withdraw the projector pole pieces, Prebus[44] inserts a secondary photographic chamber just above the projector for the specific purpose of taking diffraction patterns. Furthermore, by placing a very fine aperture in the condenser, this acting as point source for the diffraction pattern, Prebus

[42] See v. Borries and Ruska, reference 32.

[43] See O'Daniel and Radczewski, reference 33.

[44] This procedure is described in an unsigned article in *The Ohio State University Monthly*, pp. 16–18, 1943.

is able to render the illuminated area of the specimen very small, facilitating a coordination between the electron-microscopic appearance and the crystalline structure of different portions of the specimen.

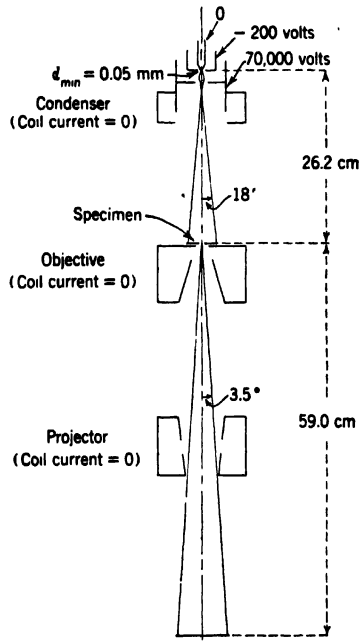


FIG. 8·23. Utilization of Siemens Electron Microscope as Diffraction Camera. (v. Borries and Ruska, reference 32. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

In the diffraction attachment of the Type B electron microscope[45] the specimen is transferred to a point just below the projector and the reduced image of the source formed by the objective is employed as point source (Fig. 8·24). The pattern is focused on the plate by a special focusing lens below the specimen. This arrangement results in a minimum convergence angle of the pencils striking the specimen and hence the greatest possible sharpness of the pattern. More important, it converts the electron microscope into an electron diffraction camera having considerable versatility. With the new specimen position it is possible to employ a holder which can rotate the specimen about any axis. In this way not only electron microscope specimens may be examined, but

[45] See Hillier, Baker, and Zworykin, reference 34.

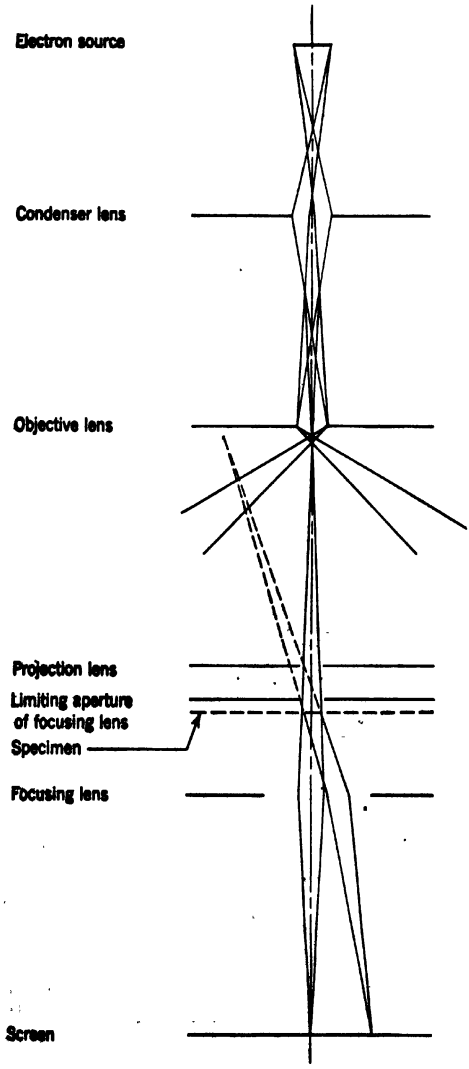


FIG. 8·24. Ray Diagram for Type B Electron Microscope Used as Diffraction Camera. (Courtesy *J. Applied Phys.*, reference 34.)

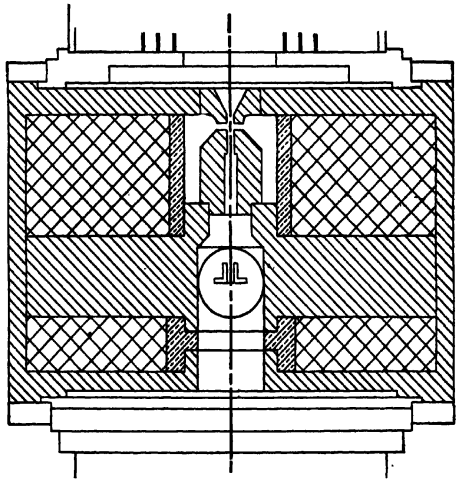


FIG. 8·25. The Diffraction Unit Replacing the Projector in the Type B Electron Microscope (Section).

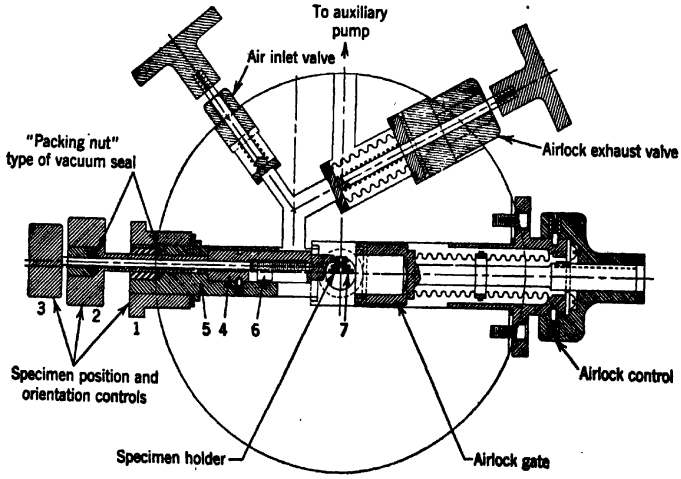


FIG. 8·26. Horizontal Section through Specimen Chamber of Experimental Diffraction Unit. (Courtesy *J. Applied Phys.*, reference 34.)

also single crystals and materials distributed on a solid backing. Such objects must be mounted so that the electron beam falls at a grazing incidence on them. Since the patterns obtained with single crystals depend very greatly on the orientation of the incident beam relative to the arrangement of the atoms in the surface plane,[46] provision must be made for rotating the specimen about an axis normal to its surface.

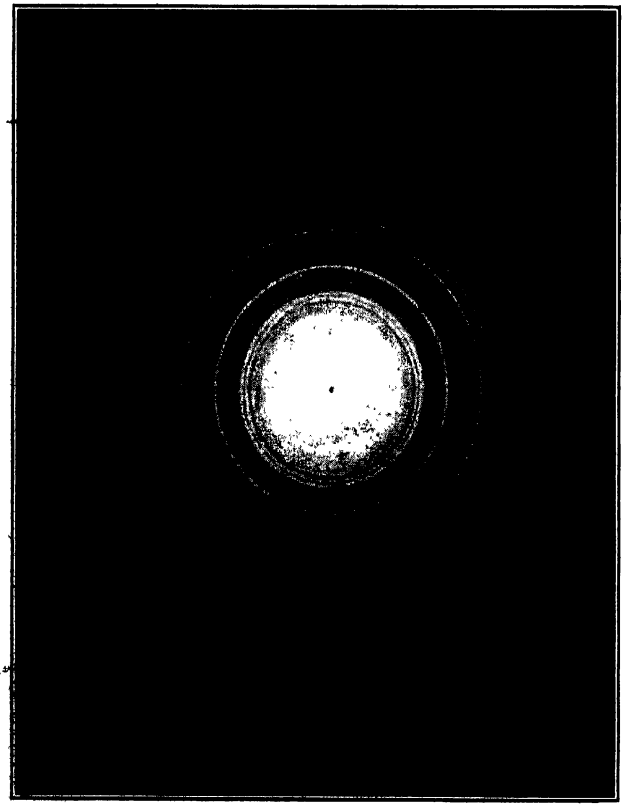


FIG. 8·27. Electron Diffraction Transmission Pattern of Zinc Smoke (ZnO).

Figure 8·25 shows a section through the diffraction unit. This unit is substituted for the projector coil assembly. The upper part is a projector lens with practically the same characteristics as the standard projector lens. Thus, in the absence of a diffraction specimen and

[46] See, for example, Meyer, reference 35, pp. 391–393.

with no current flowing through the special focusing coil at the bottom of the unit, the electron microscope with the diffraction attachment operates in the same fashion as the original instrument. However, with a specimen inserted and the stabilized projector coil current supply switched from the projector coil to the focusing coil, the micro-

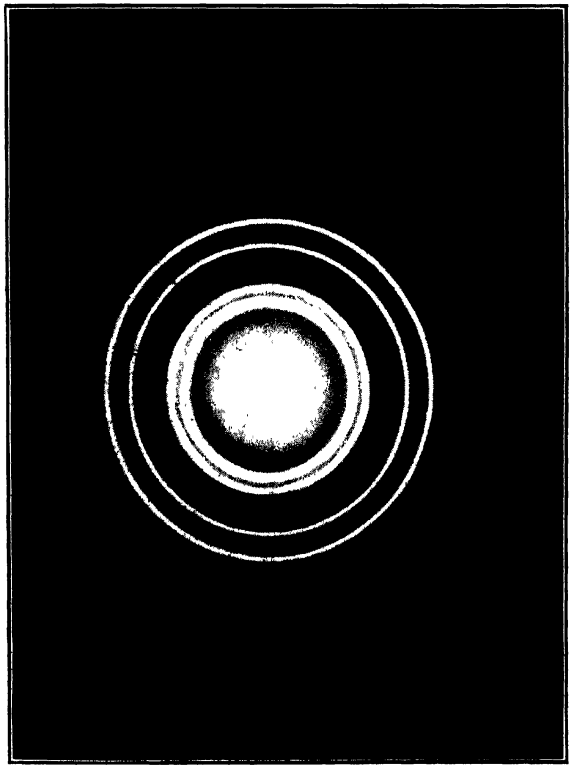


FIG. 8-28. Electron Diffraction Transmission Pattern of Evaporated Silver Film. (Courtesy *J. Applied Phys.*, reference 34.)

scope acts as a diffraction camera. A control resistor enables the operator to adjust the current in the focusing coil for the greatest sharpness. A special coil, with large-diameter pole pieces, rather than the projector lens itself, is used for focusing the beam since, for the broad beam cross sections employed, the spherical aberration of the projector lens would lead to a less sharp diffraction pattern.

A horizontal section through the specimen chamber of the camera, between the two lens coils, is shown in Fig. 8·26. The airlock arrangement is self-explanatory when compared with the specimen airlock of the electron microscope (Fig. 5·20). The holder, at the left, is designed to permit rotation of the specimen surface about an axis normal to the microscope axis, determining the angle of incidence of the beam on the specimen, and, in addition, rotation of the specimen about an axis



FIG. 8·29. Electron Diffraction Reflection Pattern of Single Crystal of Silicon.

normal to its own surface. These movements are carried out by turning milled heads 2 and 3, respectively. Turning 3 causes the nut 6 to travel along a threaded rod. As a result a flexible belt attached to it and passing around specimen table 7 rotates the table. The position of the specimen head with respect to the beam is adjusted in semipermanent fashion by fixing the position of the specimen mounting rod 4 with respect to the block 5. Transmission patterns are obtained by mounting the specimen in the usual manner on 200-mesh screen disks in a cap which fits over the specimen table and bringing the table into a horizontal position. For reflection patterns, the specimen is inserted in the hole in the table and the table is set in the position shown in Fig. 8·26.

Examples of transmission patterns obtained from a powder and from a thin metallic film are shown in Figs. 8·27 and 8·28. Figure 8·29 reproduces the reflection pattern for a silicon single crystal. The great sharp-

ness of the patterns obtained with specimens of the proper character is well illustrated by Fig. 8·27. Measurement of the rings shows that the maximum departure from a simple inverse proportionality between ring diameter and lattice spacing is 0.25 per cent. This may be ascribed to the distortion of the focusing lens rather than to the nonlinear terms in Eq. 8·7. One of the great advantages of this electron microscope diffraction camera consists in the almost perfect reproducibility of results from day to day, which make repeated calibrations with a standard substance unnecessary. Thus three sets of measurements on the patterns of three different evaporated gold films, obtained at different times in the course of a week, yielded a root-mean-square deviation of 0.03 to 0.11 per cent in the measurement of different ring diameters. This is made possible by the great constancy of the operating voltage and of the current supply for the focusing lens.

REFERENCES

1. M. v. Ardenne and H. Friedrich-Freksa, "The germination of spores of *Bacillus vulgatus* after previous imaging in the 200-kilovolt universal electron microscope," *Naturwissenschaften*, Vol. 29, pp. 523–528, August 1941.
2. F. Kirchner, "A simple method for structural examination of substances with electron waves," *Naturwissenschaften*, Vol. 18, p. 707, 1930.
3. V. J. Schaefer and D. Harker, "Surface replicas for use in the electron microscope," *J. Applied Phys.*, Vol. 13, pp. 427–433, July 1942.
4. G. Hass and H. Kehler, "A temperature-resistant and permanent supporting film for electron diffraction patterns and supermicroscopic studies," *Kolloid-Z.*, Vol. 95, pp. 29–33, March 1941.
5. M. v. Ardenne, "The object vibrator, a new accessory of supermicroscopy," *Kolloid-Z.*, Vol. 93, pp. 158–163, November 1940.
6. Columbian Carbon Company, "The particle size and shape of colloidal carbon as revealed by the electron microscope," Binney and Smith Co., distributors, New York, 1940.
7. H. Mahl, "Metallurgic studies with the electrostatic supermicroscope," *Z. tech. Physik*, Vol. 21, pp. 17–18, January 1940.
8. H. Mahl, "Supermicroscopic determination of the orientation of single aluminum crystals," *Metallwirtschaft*, Vol. 19, pp. 1082–1085, 1940.
9. H. Mahl, "Supermicroscopic surface representation by the replica process," *Naturwissenschaften*, Vol. 30, pp. 207–217, March 1942.
10. O. Rüdiger, H. Bennek, F. Stäblein, and K. E. Volk, "Structural examination of steel with the electron microscope," *Arch. Eisenhüttenw.*, Vol. 15, pp. 431–436, 1942.
11. O. Rüdiger, "The suitability of beryllium for the preparation of supermicroscopic replica films," *Naturwissenschaften*, Vol. 30, p. 279, May 1942.
12. V. K. Zworykin and E. G. Ramberg, "Surface studies with the electron microscope," *J. Applied Phys.*, Vol. 12, pp. 692–695, September 1941.
13. V. J. Schaefer, "New methods of preparing surface replicas for microscopic observation," *Phys. Rev.*, Vol. 62, pp. 495–496, November 1942.
14. R. D. Heidenreich and V. G. Peck, "Fine structure of metallic surfaces with the electron microscope," *J. Applied Phys.*, Vol. 14, pp. 23–29, January 1943.

15. M. v. ARDENNE, "On a universal electron microscope for bright-field, dark-field, and stereo operation," *Z. Physik*, Vol. 115, pp. 339–368, March 1940.
16. M. v. ARDENNE, "Electron microscopy of living substance," *Naturwissenschaften*, Vol. 29, pp. 521–523, August 1941.
17. B. v. BORRIES and E. RUSKA, "A supermicroscope for research laboratories," *Naturwissenschaften*, Vol. 27, pp. 577–582, 1939.
18. C. J. BURTON, R. B. BARNES, and T. G. ROCHOW, "The electron microscope, calibration and use at low magnifications," *Ind. Eng. Chem., Ind. Ed.*, Vol. 34, pp. 1429–1436, December 1942.
19. W. BOTHE, "Photographic β-ray measurements," *Z. Physik*, Vol. 8, pp. 243–250, 1922.
20. M. v. ARDENNE, "Resolving power of photographic emulsions for electron radiations," *Z. Physik*, Vol. 114, pp. 379–388, November 1939.
21. B. v. BORRIES, "Intensity relations in the supermicroscope: I. Blackening of photoemulsions by electron beams," *Physik. Z.*, Vol. 43, pp. 190–204, 1942.
22. B. v. BORRIES, "Intensity relations in the supermicroscope: II. Capacity of enlargement, graininess, and resolving power of photoplates blackened by electrons," *Z. angew. Photogr.*, Vol. 4, pp. 42 ff., 1942.
23. B. v. BORRIES, "Intensity relations in the supermicroscope: III. Suitableness and limits of sensitivity of photographic plates for supermicrography," *Z. Physik*, Vol. 119, pp. 498–521, 1942.
24. M. v. ARDENNE, "On the usefulness of photoemulsions with little binder for electron supermicroscopy," *Z. Physik*, Vol. 121, pp. 1–6, 1943.
25. R. F. BAKER, E. G. RAMBERG, and J. HILLIER, "The photographic action of electrons in the range between 40 and 212 kilovolts," *J. Applied Phys.*, Vol. 13, pp. 450–456, July 1942; Vol. 14, p. 39, January 1943.
26. W. HENNEBERG, "The supermicroscope with electrostatic lenses," *Elektrotech. Z.*, Vol. 61, pp. 773–776, 1940.
27. L. MARTON and L. I. SCHIFF, "Determination of object thickness in electron microscopy," *J. Applied Phys.*, Vol. 12, pp. 759–765, October 1941.
28. W. EITEL and E. GOTTHARDT, "The stereophotogrammetric measurement of the thickness of very small crystals from supermicroscopic pictures," *Naturwissenschaften*, Vol. 28, p. 367, June 1940.
29. E. RUSKA and H. O. MÜLLER, "Technical improvements in the formation of images of surfaces irradiated with electrons," *Z. Physik*, Vol. 116, pp. 366–369, 1940.
30. B. v. BORRIES, "Higher resolving power in the formation of images of surfaces by the supermicroscope," *Z. Physik*, Vol. 116, pp. 370–378, 1940.
31. J. D. HANAWALT, H. W. RINN, and L. K. FREVEL, "Chemical analysis by x-ray diffraction," *Ind. Eng. Chem., Anal. Ed.*, Vol. 10, pp. 457–512, September 1938.
32. B. v. BORRIES and E. RUSKA, "High-resolution microscopy with fast electrons," *Ergebn. d. exakt. Naturwiss.*, Vol. 19, pp. 237–322, 1940.
33. H. O'DANIEL and O. E. RADCZEWSKI, "Electron microscopy and diffraction of highly disperse minerals on the same specimen," *Naturwissenschaften*, Vol. 28, pp. 397–399, June 1940.
34. J. HILLIER, R. F. BAKER, and V. K. ZWORYKIN, "A diffraction adapter for the electron microscope," *J. Applied Phys.*, Vol. 13, pp. 571–577, September 1942.
35. C. F. MEYER, *The Diffraction of Light, X-Rays, and Material Particles*, University of Chicago Press, Chicago, 1934.

CHAPTER 9

THE ELECTRON MICROSCOPE AS A RESEARCH INSTRUMENT

9·1. Present Range of Application of the Electron Microscope. The electron microscope has been known for over a decade and its capabilities as a research instrument have been recognized for a considerable fraction of this period. At first sight, therefore, it may seem somewhat surprising that a vast literature of results and achievements is not already in existence. This becomes understandable, however, when it is recognized that it is only within the past three or four years that the instrument has been made sufficiently practical to become a scientific tool for the general research laboratory. Now, with the instrument more generally available, the literature is building up rapidly, and every year brings an increasing number of important contributions in many fields of science.

This chapter cannot attempt to cover all or even the major part of the various contributions made to scientific research with the aid of the electron microscope. However, a number of investigations in various fields will be outlined, indicating the ability of this instrument to gather information inaccessible to the light microscope and to other means of direct observation. The examples chosen do not necessarily represent the outstanding or most recent work in any given line but are selected merely as typical of what can be accomplished. To acquire a more complete grasp of the results which have been achieved the reader is referred to the many excellent original articles.

The greatest variety of problems to which the electron microscope has been applied is encountered in the field of biology. This work ranges from investigations of animal tissue, by methods only recently worked out, through bacteriology and bacterial morphology to studies of viruses, bacteriophage, and other minute biological entities beyond the limits of the light microscope. However, neither the potentialities nor the work already done by the electron microscope is limited to this field. Thus the new instrument has found application in both scientific and industrial physics and chemistry. It has been utilized in initial investigations of plastics, studies of carbon black and its effect on rubber, work with colloidal particles, and observations of crystals and crystal

surfaces. More recently it has found wide application in the routine determination of particle size and structure of all types of bulk materials. This experience suggests that the electron microscope may contribute as much or more in the field of chemistry as in the field of biology, bridging, as it does, most of the gap between structures that can be investigated by means of x-ray and electron diffraction and those which may be observed directly with the light microscope.

The usefulness of the electron microscope in the study of metals has been greatly extended by methods of preparing surface replicas of metallographic sections described in section 8·1. Studies of several metals and alloys have recently been made in this manner. Although this type of research is as yet in its infancy, it is safe to predict that the electron microscope will prove to be a powerful tool in the hands of the metallurgist.

9·2. Biological Research. Chronologically the work in the field of biology was begun with an investigation of methods of studying bacteria and organic cell tissue. The initial work, done by L. Marton in 1934,[1] involved saturating the specimen with an osmium salt, which, after the organic material had been destroyed, left a fossil of the original organic structure. This fossil was micrographed. Shortly after this it was found that specimens such as bacteria supported on a collodion film could be observed and micrographed in an electron microscope without destruction by the electron beam.[2] This method, improved and refined in certain details, is employed at present. The material to be examined is deposited upon a thin film of collodion or a similar plastic. The supporting film is mounted on a fine mesh or stretched across an aperture in a small disk and placed in a holder which allows it to be inserted into the microscope. Details of the formation and manipulation of the supporting film have already been given in section 8·1.

The electron microscope has proved particularly effective in the study of viruses. Before its advent most viruses were beyond the range of direct observation. It is, therefore, not strange that results in the field of virology obtained with this new tool were awaited with great interest.

Viruses are minute particles believed to consist in part or wholly of nucleoproteins. They are known to biologists as obligate intracellular parasites; in other words, they do not multiply except in the cells of the proper living host. All attempts to grow cultures of these organisms in artificial media have met with failure. Viruses are not, however, necessarily destroyed when removed from their host. It is hence possible to isolate and purify them by suitable techniques, such as ultracentrifuga-

[1] See reference 1.

[2] See Marton, reference 2.

tion and ultrafiltration. When reintroduced into a living host these purified strains of virus will resume their activity. Various indirect methods had shown their sizes to range between 10 and 250 millimicrons. These methods had, furthermore, indicated that both spherical and asymmetrical viruses existed. However, little exact information was available as to their size, size distribution, and shape, and none at all concerning their structure.



FIG. 9·1. Tobacco Mosaic Virus from a Dilute Suspension. (Courtesy *J. Biol. Chem.*, reference 3.)

The first work on viruses with the electron microscope had the character of orientation studies. An investigation of this type is illustrated by the work of W. M. Stanley of the Rockefeller Institute and T. F. Anderson, RCA National Research Council Fellow, on purified plant viruses.[3] Tobacco mosaic virus, rather thoroughly studied beforehand by other methods, was selected as one of the first subjects for examination. The virus was purified by differential centrifuging. Specimens of different concentrations were placed on film supports and allowed to dry. Figure 9·1[4] is a micrograph made of a very dilute suspension. The minute individual rodlike particles are readily distinguished. They have an average length of 280 millimicrons and a width of 15 millimi-

[3] See Stanley and Anderson, reference 3. For earlier work in the same field, see Kausche, Pfankuch, and Ruska, reference 4.

[4] The magnification of the micrograph is indicated throughout by a marker showing the apparent length of a one-micron (1 μ) separation in the object. For example, if this marker is 1 cm in length the magnification is 10,000. For micrographs with very high and with relatively low magnification, markers for 0.1 μ and 10 μ, respectively, are substituted.

crons. Measurements on a large number of these entities show that there is some deviation from these dimensions; yet, this deviation is small. When the concentration of particles is high and when the preparation has aged, a tendency for end-to-end and side-to-side attachment is exhibited, which is shown in Fig. 9·2. It may give the specimen an almost crystal-like appearance.

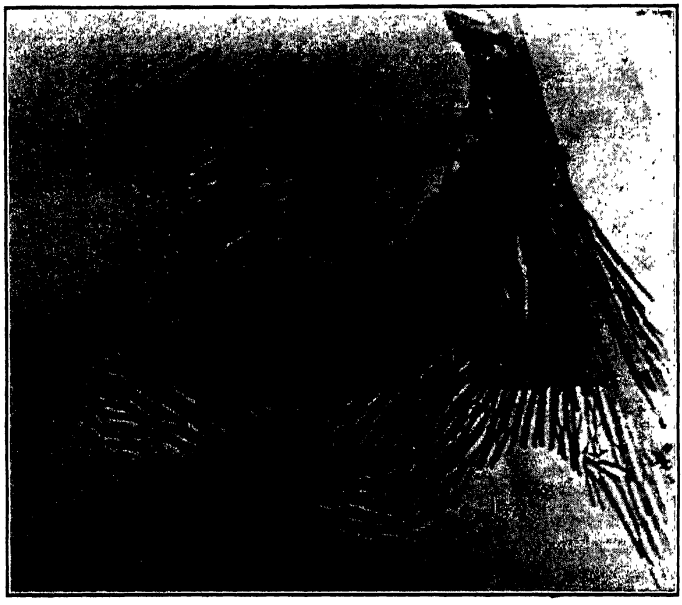


FIG. 9·2. Aggregation of Tobacco Mosaic Virus Particles. (Courtesy *J. Biol. Chem.*, reference 3.)

Similar to tobacco-mosaic virus in their rodlike shape and their tendency to form end-to-end aggregates, cucumber-mosaic particles (strains 3 and 4) were found to have a length of about 300 millimicrons. Whereas tobacco mosaic virus thrives in a large number of only remotely related plant hosts, the activity of cucumber mosaic virus is restricted to the family of the Cucurbitaceae.

By no means all plant virus particles are rodlike in shape. Thus tomato bushy stunt virus and tobacco necrosis virus, shown in Figs. 9·3*a* and 9·3*b*, consist of approximately spherical elements.

Ever since viruses were first identified there have been speculations as to whether the virus particles might be regarded as undifferentiated

macromolecules. Indeed, biologists have frequently designated such particles as virus "molecules." In the case of certain of the viruses, however, the electron microscope has been able to prove conclusively that the individual particles possess relatively elaborate structural differentiations.

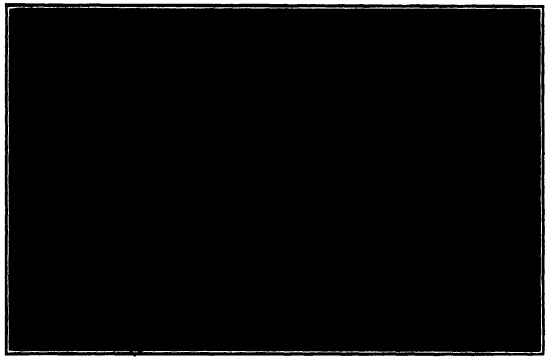


FIG. 9·3(a). Tomato Bushy Stunt Virus. (Courtesy *J. Biol. Chem.*, reference 3.)

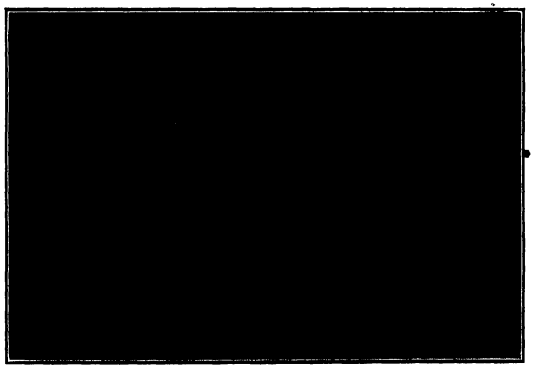


FIG. 9·3(b). Tobacco Necrosis Virus. (Courtesy *J. Biol. Chem.*, reference 3.)

The first evidence of such structure in viruses was found in the study of bacteriophage. Bacteriophage constitutes a group of viruses which reproduce in the presence of living bacterial cells and destroy their hosts in the process. Inasmuch as certain of them prey on pathogenic bacteria they are of interest to medical research.

Early electron microscope observations of bacteriophage by Pfankuch

and Kausche[5] indicated merely that the phage particles were minute round bodies. It remained for Luria and Anderson[6] to demonstrate not only that the phage particles of the strain anticoli PC were tadpolelike in shape, with a clearly differentiated head and tail, but, furthermore, that the head of the phage particle contained a characteristic arrange-

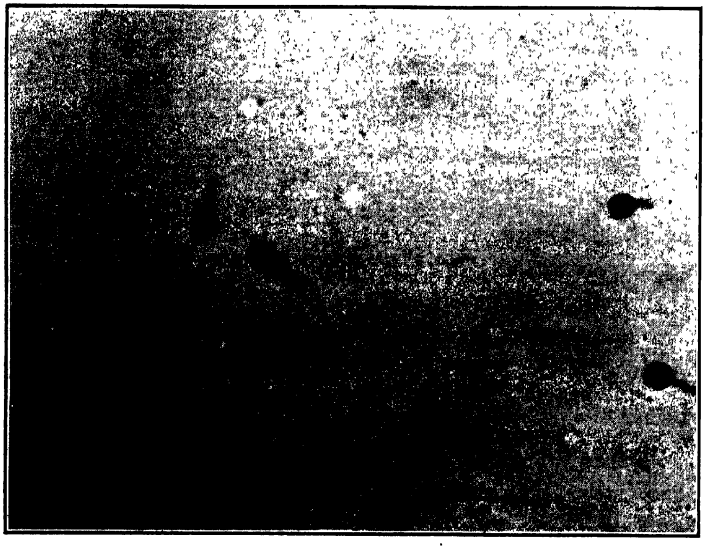


FIG. 9·4. Isolated Bacteriophage Anticoli PC. (Courtesy *Proc. Natl. Acad. Sci. U.S.*, reference 6.)

ment of scattering material. These markings are clearly visible in Fig. 9·4. Other micrographs show how the phage particles multiply within the host cell (*Escherichia coli*) until the cell wall bursts with explosive violence, scattering phage particles and remnants of the cell content (Fig. 9·5).

Another virus which, when examined with the electron microscope, clearly reveals internal structure is the virus of vaccinia or cowpox.[7] Figure 9·6 shows the elementary bodies of vaccinia from a purified preparation. The arrangement of the darker, or denser, areas within the rectangular-shaped virus particles is indicated in the drawings below them. These particles, which are about 250 millimicrons in diameter,

[5] See reference 5.
[6] See reference 6.
[7] See Green, Anderson, and Smadel, reference 8.

are large enough to be examined with the light microscope when stained. Such an examination, however, yields no information regarding shape or structure.

The control and destruction of viruses present an important problem both in medicine and in plant breeding. They may be destroyed by adding a chemical, such as a small amount of ammonia, to the virus

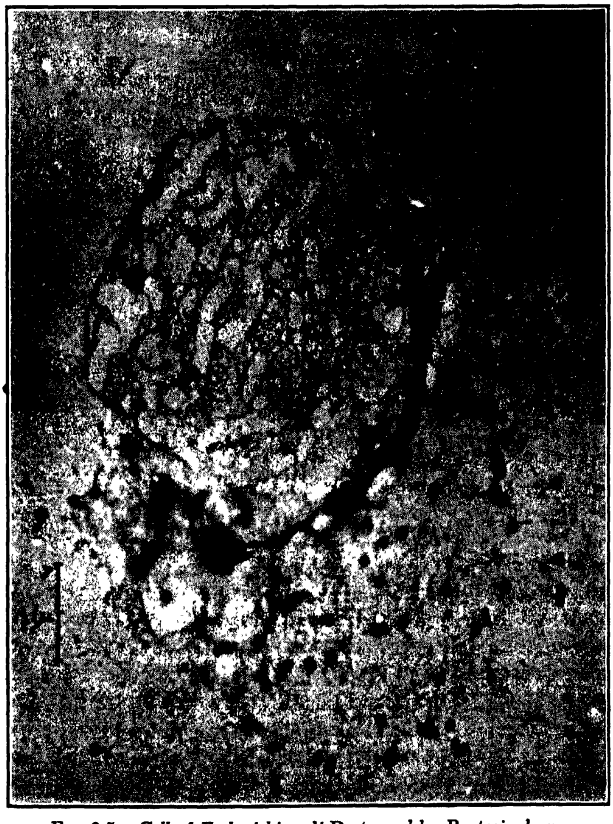


FIG. 9·5. Cell of *Escherichia coli* Destroyed by Bacteriophage. (Courtesy *J. Bact.*, reference 7.)

suspension. Viruses are also inactivated by the addition of serum from animals into whose blood stream the same virus strain has previously been injected. Anderson and Stanley[8] have studied a number of

[8] See reference 9.

inactivating reactions of this character with the electron microscope. The animal body, under these circumstances, generates *antibodies* to protect itself against the invaders. The action of these antibodies is highly specific. Thus, if serum from a rabbit injected with tobacco mosaic virus is added to a mixed suspension of tobacco mosaic and

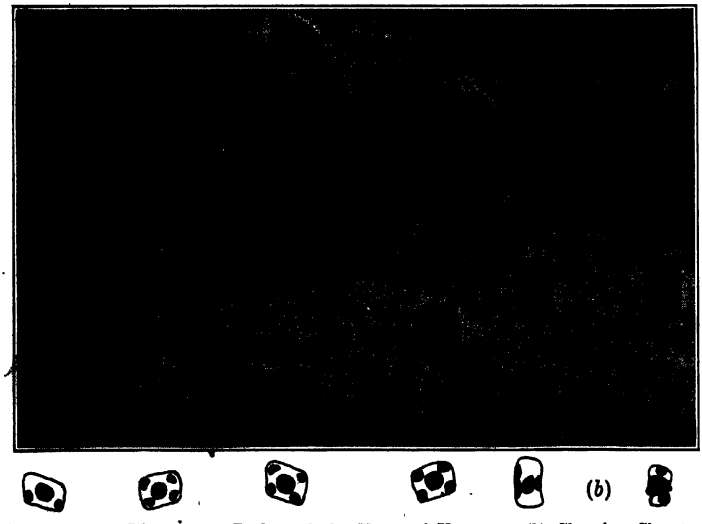


FIG. 9·6. (*a*) Elementary Bodies of the Virus of Vaccinia; (*b*) Sketches Showing Granule Arrangement in Elementary Bodies. (Courtesy *J. Exptl. Med.*, reference 8.)

tomato bushy stunt virus, only the tobacco mosaic virus will be inactivated, the bushy stunt virus remaining unaffected (Fig. 9·7*a*). Conversely, serum from a rabbit injected with bushy stunt virus added to the same mixture will affect the bushy stunt virus alone (Fig. 9·7*b*). The inactivation of the virus particles reveals itself in a considerable broadening and a tendency to clump together. The broadening is believed to be due to the endwise attachment of the antibodies, which, to the best of our knowledge, are elongated molecules approximately 17 millimicrons long and 4 millimicrons wide. The clumping may be caused by the reactivity of the outer ends of the antibody molecules or by some other colloidal effect.

Other studies have concerned themselves with the presence of tobacco mosaic virus in the chloroplasts of diseased plants[9] and with its inactivation by substances liberated in tobacco seed during germination.[10]

[9] See Kausche and Ruska, reference 10.
[10] See Kausche, reference 11.

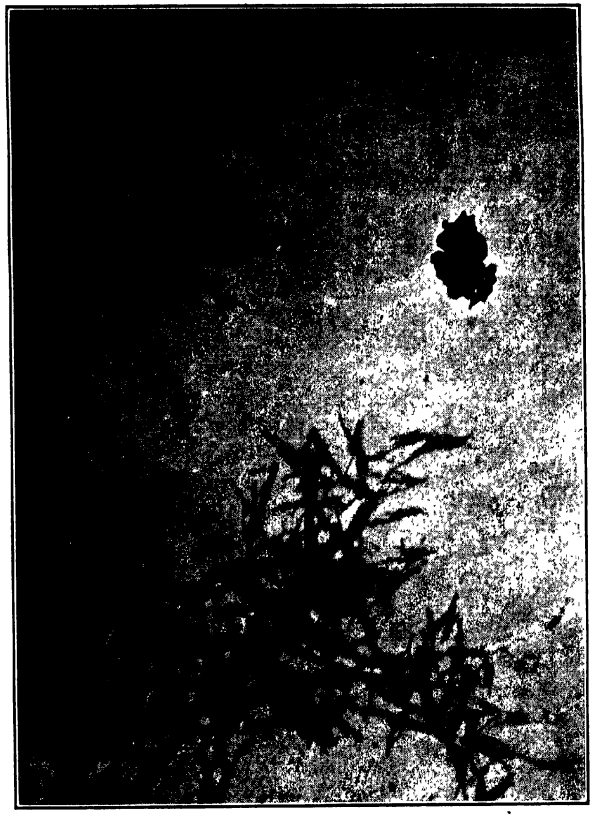


FIG. 9·7(*a*). The Effect on a Mixed Suspension of Tobacco Mosaic Virus and Tomato Bushy Stunt Virus of Adding Antitobacco Mosaic Serum. (Courtesy *J. Biol. Chem.*, reference 9.)

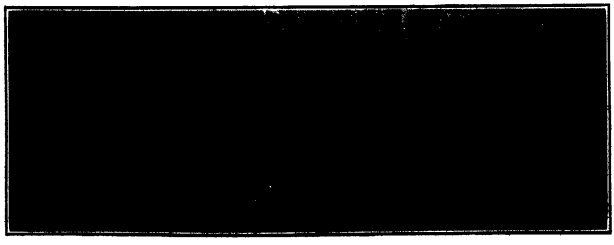


FIG. 9·7(*b*). The Effect on a Mixed Suspension of Tobacco Mosaic Virus and Tomato Bushy Stunt Virus of Adding Antibushy Stunt Serum. (Courtesy *J. Biol. Chem.*, reference 9.)

Successful attempts have also been made to render visible the causative agents of various animal virus diseases[11] other than vaccinia.

The function of the electron microscope in the study of bacteria differs somewhat from that in virus research. These organisms are large enough not only to be visible in the light microscope, but even to

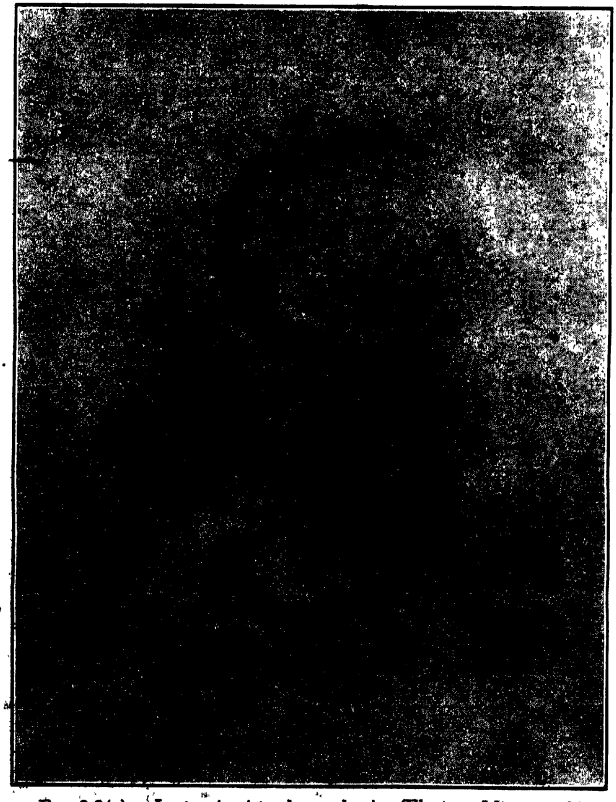


FIG. 9-8(a). *Leptospira icterohemorrhagiae* (Electron Micrograph).

reveal their general shape. Their familiar subdivision into cocci (spherical), bacilli (rod-shaped), and spirillae (corkscrew-shaped bacteria) is based on this circumstance. In some bacteria the presence of internal structure — in the form, for example, of polar granules — has

[11] See v. Ardenne and Pyl, reference 12, v. Ardenne, reference 13, p. 361, and Sharp, Taylor, Beard, and Beard, references 14 and 15.

even been demonstrated. The service to be performed by the electron microscope must, therefore, consist primarily in a more detailed analysis of the structure of bacteria and the role which different parts play in bacterial reactions. Furthermore, the more intimate view obtained in

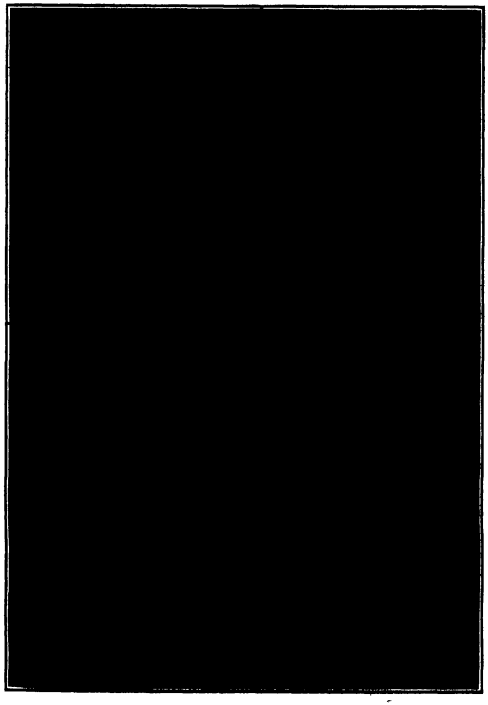


FIG. 9·8(*b*). *Bacillus subtilis* (Electron Micrograph).

this manner may facilitate the differentiation of different strains, varying in their physiological effect, of a particular bacterium. Figures 9·8 and 9·9, comparing electron micrographs and light micrographs of three familiar types of bacteria — a spirochete, *Bacillus subtilis*, and a typhoid bacterium — give some indication of the promise of the new method. The long flagella visible on the electron micrographs may be observed in the light microscope only with the aid of a staining technique which completely distorts their natural appearance.

Detailed morphological studies of *Streptococcus pyogenes* have been carried out by Mudd and Lackman[12] and of *Bacillus subtilis*, *B. anthracis*,

[12] See reference 16.

and *B. megatherium* by Mudd, Polevitzky, Anderson, and Chambers.[13] All these bacteria were found to be enveloped by strong cell walls, which were not distorted by introduction into vacuum. In the streptococcus this cell wall tended to extend continuously over long chains of cocci, holding them together (Fig. 9·10). At points of future division the cell wall is indented and an intercellular plate develops, as is brought

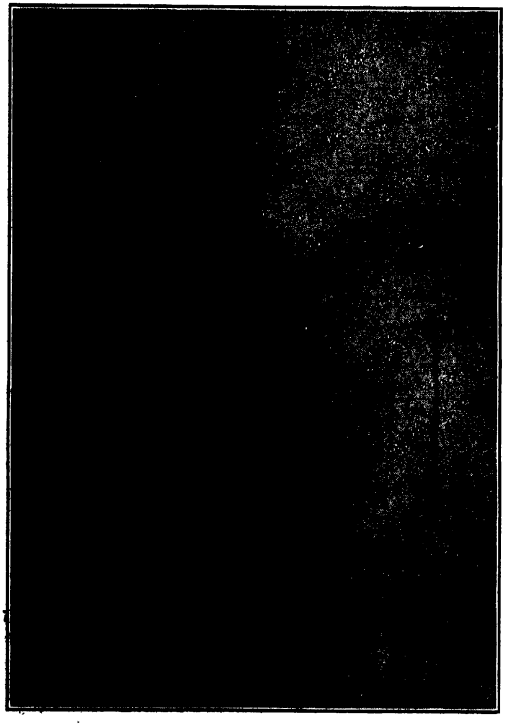
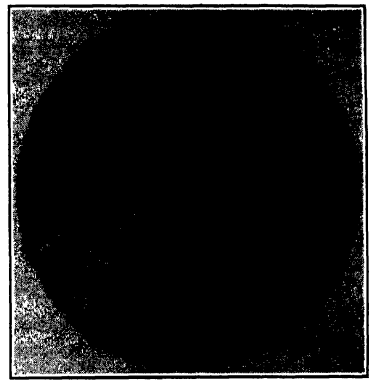


FIG. 9·8(*c*). *Eberthella typhosa* (Electron Micrograph).

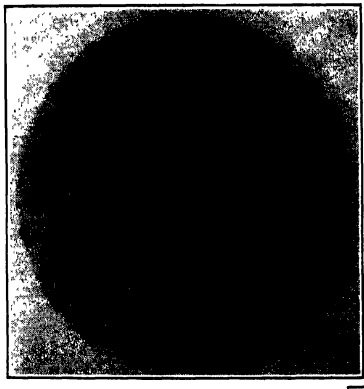
out clearly in the picture of *Bacillus anthracis* (Fig. 9·11). If the cells are fragmented by sonic vibrations, the protoplasm escapes, leaving behind the empty cell walls. In old cells the protoplasm shrinks from the cell membrane. Here the continuity of the flagella with the cell membrane becomes clearly evident (Fig. 9·12).

Further electron microscope studies of bacterial structure were carried out, particularly for a large number of anaerobic bacteria, by

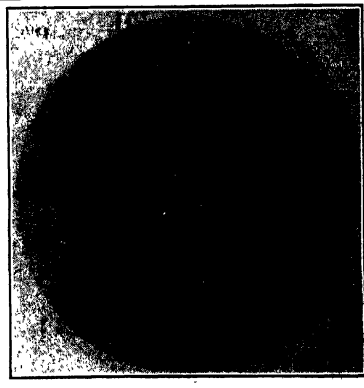
[13] See reference 17.



(a). *Leptospira icterohemorrhagiae.*



(b). *Bacillus subtilis.*



(c). *Eberthella typhosa.*

FIG. 9·9. Light Micrographs of Familiar Types of Bacteria. (Courtesy Dr. H. E. Morton, University of Pennsylvania.)

Jakob and Mahl.[14] These authors held that the cell wall constituted a capsule from which the bacterial body might, under favorable circumstances, emerge, an interpretation held untenable by Mudd and coworkers. Lembke and Ruska[15] have made a detailed examination of

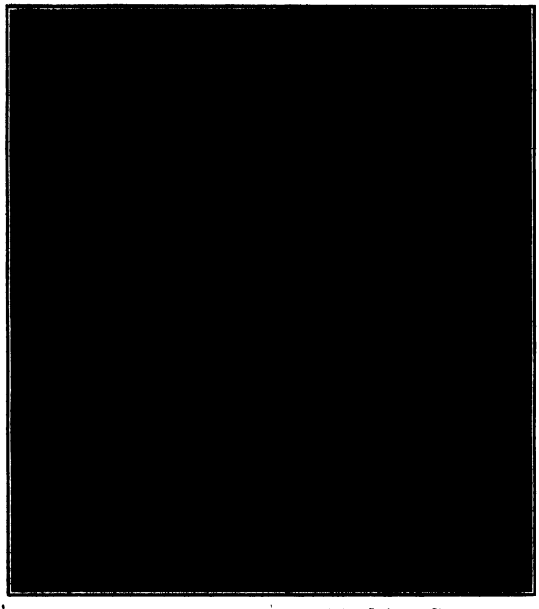


FIG. 9·10. Streptococcus from a Blood Agar Culture.
(Courtesy *J. Bact.*, reference 16.)

the internal structure of the tubercle bacillus, the avian type in particular. They found, in addition to the familiar dense granules, much smaller *microgranules* and regions of low density, or *vacuoles*. Microgranules appeared to be formed within the vacuoles as the result of a separating-out process, eventually filling them up completely and thus forming the larger granules. A further detailed study of a single bacterium, a particular strain of *Treponema pallidum*, the causative germ of syphilis, has been published by Morton and Anderson.[16] Numerous dense granules are observed within the bacterial body, protruding from it, or attached to it in clusters. Constrictions, probably pre-

[14] See reference 18.
[15] See reference 19.
[16] See reference 20.

liminary to a cell division, are often observed. Finally, flagellar filaments are found attached to the cell at various points. A number of these properties are observable in Fig. 9·13.



FIG. 9·11. *Bacillus anthracis*, Showing Thickened Cell Membrane between Adjoining Cells. (Courtesy *J. Bact.*, reference 17.)

Other research has been concerned with bacterial reactions rather than with normal cells. Thus, it had been known for some time that diphtheria bacilli, bred on a nutrient containing potassium tellurite, would reduce the tellurite, precipitating metallic tellurium. This gave the culture a dark coloration. It was not known, however, whether this reduction took place inside, or at the surface of, the bacterial cells.

The question was settled by Morton and Anderson[17] with the aid of the electron microscope. As shown in Fig. 9·14, diphtheria bacilli grown on tellurite chocolate agar reveal, in their interior, needle-shaped tellurium crystals which, in several cases, distort the cell walls. The tellurite salt must thus diffuse through the wall before reduction. An addition of bromine water to the culture causes the needles to vanish, forming a soluble tellurium salt.

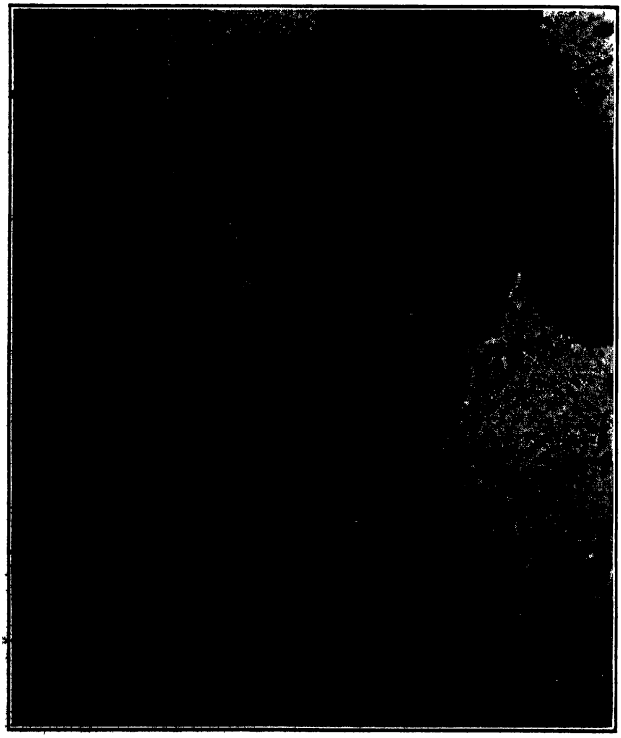


FIG. 9·12. Spores of *Bacillus subtilis* with Protoplasm Shrunken from Cell Membrane. (Courtesy *J. Bact.*, reference 17.)

Staining reactions of fusobacteria as well as of typhoid, dysentery, and cholera bacteria treated with heavy-metal salts have been investigated by Mudd and Anderson.[18] To obtain the specimen, a droplet of

[17] See reference 21.
[18] See reference 22.

the bacterial suspension is deposited on a collodion film. A drop of heavy-metal solution from a micropipette is permitted to fuse with it. After time has been allowed for the reaction to take place, the liquid is removed by placing it in contact with the meniscus of distilled water in

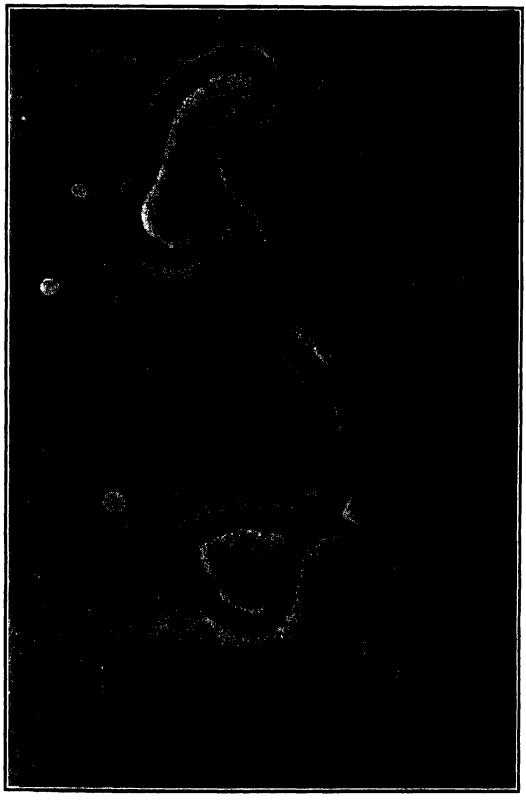


FIG. 9·13. *Treponema pallidum*, Showing Flagella and Granules.

a test tube held nearly horizontal. The bacteria which have come to adhere to the collodion film are not disturbed by the process, which may be employed to wash the specimen repeatedly. Some of the results obtained in this manner with typhoid bacteria are shown in Fig. 9·15. It is seen (Fig. 9·15*b*) that the addition of silver nitrate results in a contraction of the cells, the destruction of the flagella, and a dense staining of the protoplasmic cell content. Cells treated with lead acetate appear

swollen and surrounded by a halo of escaped protoplasm (Fig. 9·15*c*). The staining effect of the lead salt is confined to the cell membranes and the flagella. The flagella appear darker without being thickened. In the regions near the bacteria the accumulation of protoplasmic components against the flagella during drying creates the illusion of a tubular structure.

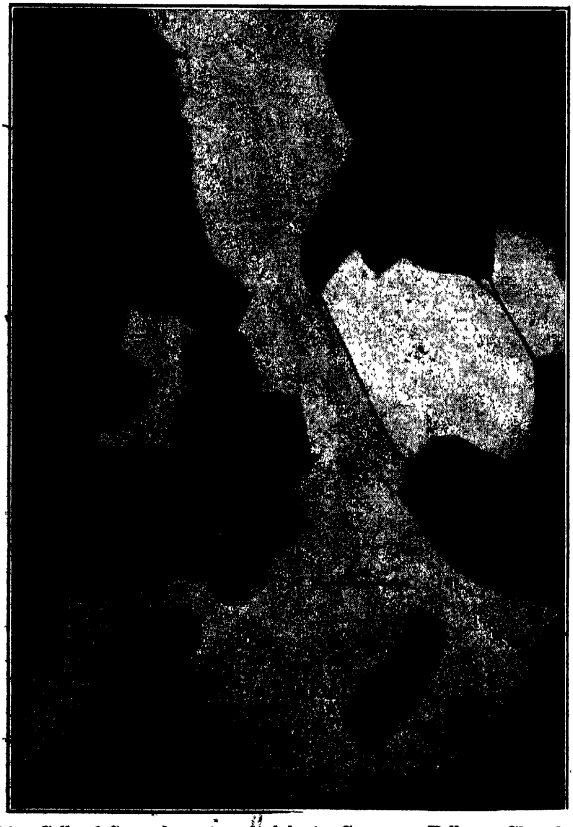


FIG. 9·14. Cells of *Corynebacterium diphtheriae* Grown on Tellurite Chocolate Agar; Needle-Shaped Tellurium Crystals Are Visible within the Cells. (Courtesy *Proc. Soc. Exptl. Biol. Med.*, reference 21.)

The same authors[19] have made a study of the reaction of typhoid, paratyphoid, and subtilis bacilli with rabbit antisera, obtained from

[19] See Mudd and Anderson, reference 23.

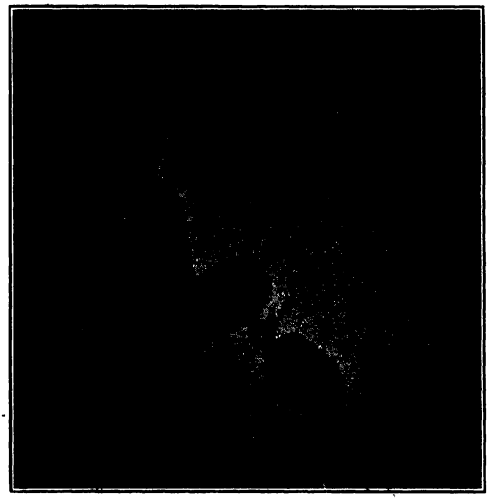


FIG. 9·15(a). Cells of *Eberthella typhosa* Dried from Distilled Water. (Courtesy *J. Exptl. Med.*, reference 22.)

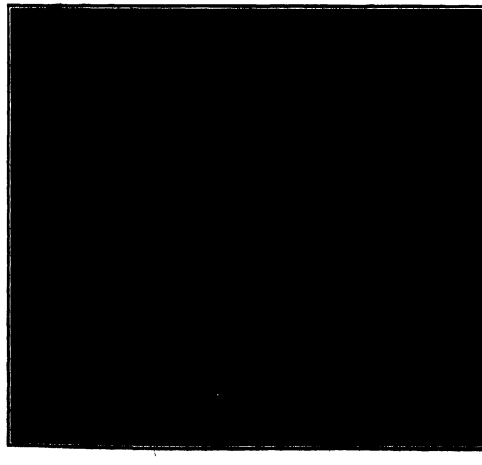


FIG. 9·15(b). Cells of *Eberthella typhosa* after Exposure to 1.5 Molar $AgNO_3$ Solution. (Courtesy *J. Exptl. Med.*, reference 22.)

animals injected with various bacterial cultures. As in the case of the corresponding virus reactions, heterologous antisera are found to have no effect on the bacteria (Fig. 9·16*a*). However, the addition of the homologous serum — for example, of antityphoid serum to *Eberthella typhosa* (Fig. 9·16*b*) — causes a thickening of the flagella and, in addi-

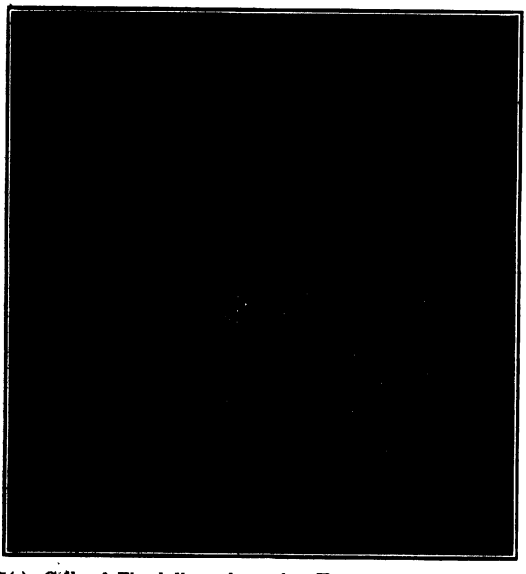


FIG. 9·15(*c*). Cells of *Eberthella typhosa* after Exposure to 0.29 Molar $Pb(Ac)_2$ Solution. (Courtesy *J. Exptl. Med.*, reference 22.)

tion, a deposition of antibody protein molecules on the surface of the cell wall. The broadening of the flagella is considerably less than the broadening observed in the case of the viruses (Fig. 9·7). This may be ascribed to the fact that, with the low concentrations of antiserum here employed, the elongated antibody molecules may attach themselves lengthwise or at a small angle to the flagella rather than radially. This explanation, however, is only tentative.

In the study of the higher animals, the examination of the constituents of the blood stream assumes an intermediate position. Here, since these entities are of magnitude comparable to those of bacteria, the methods of preparation resemble those familiar from bacteriology. A study of the degeneration of red corpuscles (erythrocytes) under the influence of certain poisons has been made by Jung.[20] Of more con-

[20] See reference 24.

venient dimension for electron-microscopic investigation are the blood plates and fibrin, two constituents playing the leading role in the process of blood clotting. Their morphology and the transformation of the blood plates during the clotting process have been studied in great

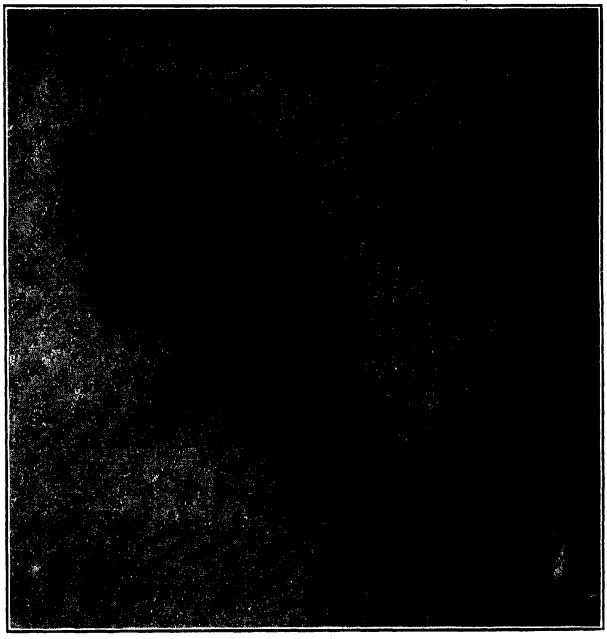


FIG. 9·16(*a*). *Eberthella typhosa* with Antistreptococca Serum. (Courtesy *J. Immunol.*, reference 23.)

detail by Wolpers and Ruska.[21] The granules liberated by the disintegrating blood plates act as centers of cohesion for the fibrin structure of the blood clot.

Whenever the material to be micrographed is a fragment of a larger organism or a section of tissue the technique of preparation becomes much more difficult. Methods have to be evolved for subdividing, cutting, or sectioning the material to a fineness otherwise not required and for transferring the specimen to the supporting film without injury. A good illustration of such techniques is given in an investigation by Richards and Anderson of the cuticle of the American cockroach and of the larva of the mosquito.[22]

[21] See reference 25.
[22] See reference 26.

It soon became apparent that the ordinary sectioning methods could not be applied to obtain a transverse section of the cockroach cuticle thin enough to be observed with the electron microscope. Even tapered microtome sections did not prove satisfactory, since the thin end would

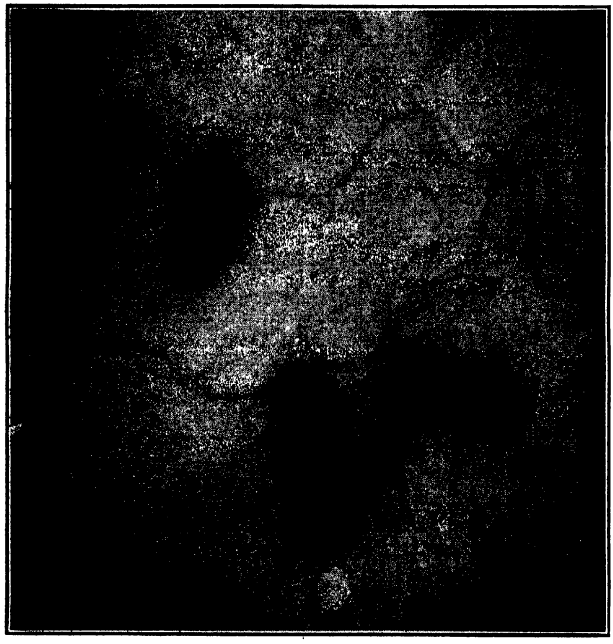


FIG. 9·16(*b*). *Eberthella typhosa* with Antityphoid Serum. (Courtesy *J. Immunol.*, reference 23.)

not hold together mechanically. After several trials the following technique was adopted. A fresh fragment of cuticle cleaned by swabbing away all extraneous tissue was placed in a slit cut into a turgid carrot. After the slit had been clamped with a screw clamp, a level cutting surface was prepared with a single razor cut. A sharp razor blade was then slid back and forth across the cutting surface, with only slight pressure. In this manner thin sections were sliced off as the surface of the carrot was gradually worn down. This material was washed off into a watch glass. Periodically, the thinner sections were selected from the glass with the aid of a binocular microscope and transferred to a second dish containing distilled water to which a slight amount (about 0.2 per cent) of albumen fixative had been added. After further sorting

teristic spines, which had not been observed with the light microscope. These spines are absent in the tracheal branches. Even the smallest breathing tubes, the tracheoles, however, were found to possess tenidiae. In certain sections of the tracheal trunk the tenidiae are missing, the membrane being instead provided with large spines (Fig. 9·18). Stereomicrographs were employed to advantage, indicating, in some cases, one tracheal tube within the other, thus proving that the complete tracheal system had not been removed in the previous moult.

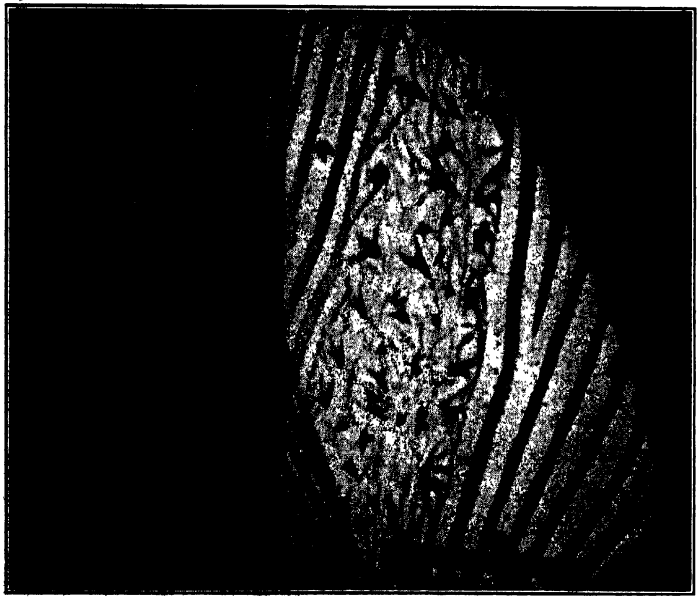


FIG. 9·18. Portion of Spinose Band of Longitudinal Tracheal Trunk of Mosquito Larva. Accelerating Potential: 200 Kilovolts. (Courtesy *J.N.Y. Entomol. Soc.*, reference 27.)

The same stereotechnique was employed by Anderson and Richards[24] to interpret the optical mechanism giving rise to various insect colorations. Figure 9·19 shows an iridescent scale of the butterfly *Morpho cypris*. The coloration is here due to interference between the light reflected from successive layers of structural elements.

Morphological studies have also been made of chromosomes, muscle

[24] See reference 28.

tissue,[25] the fibrillar structure of the nerve axoplasm of the squid,[26] and collagen, the chief constituent of the fiber of connective tissue and the organic matter of bones.[27] In many investigations of this type the treatment of the specimen with selective electron stains of high mass and molecular weight is found advantageous, increasing the contrast in the electron image. Figure 9·20 shows, as an example, an electron



FIG. 9·19. Scale of Butterfly Morpho Cypris — Stereomicrographs. (Courtesy *J. Applied Phys.*, reference 28.)

micrograph, obtained with the electron microscope at the Massachusetts Institute of Technology, of typical fibrils from the adductor muscle of the clam *Venus mercenaria* stained with phosphotungstic acid.[28a] The stain combines with the protein at regular intervals along the fiber axis as a result of a periodicity in the protein structure. The average spacing as determined from some 200 fibrils was found to be 145 A.U., with a maximum deviation from this value of about 10 per cent.

If the stain is applied in suitable amount, each cross band can be seen to consist of a series of equally spaced spots, as shown in Fig. 9·21*a*.

[25] See Richards, Anderson, and Hance, reference 29.

[26] See Richards, Steinbach, and Anderson, reference 30.

[27] See Hall, Jakus, and Schmitt, reference 31*a*.

[28a] This work was reported by C. E. Hall, M. A. Jakus, and F. O. Schmitt at the Chicago meeting of the Electron Microscope Society of America on November 16, 1944.

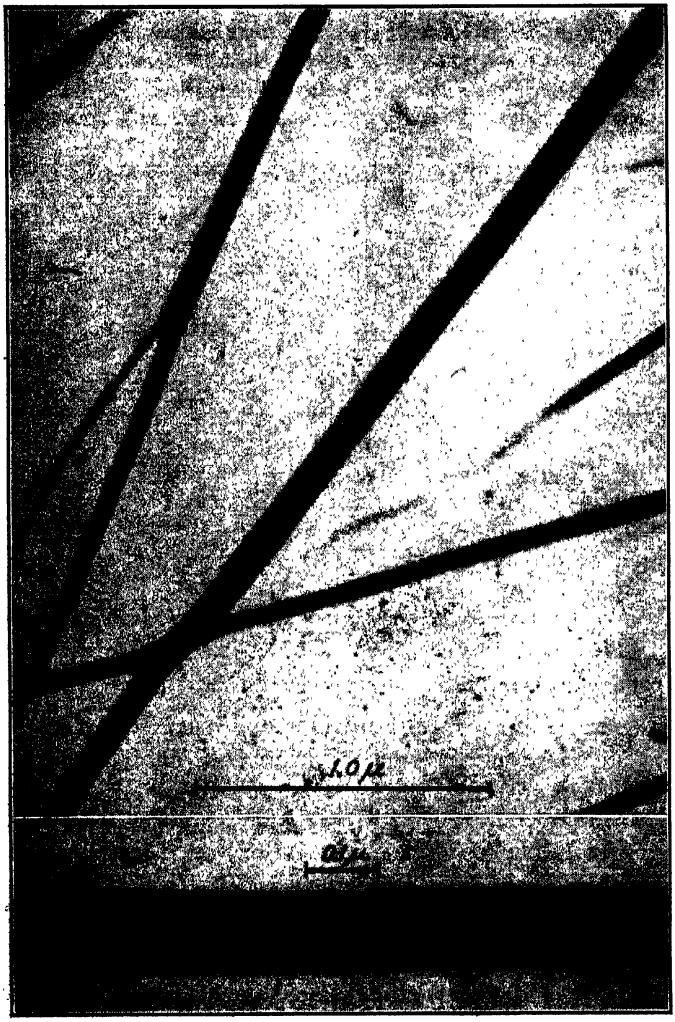


FIG. 9·20. Fibrils from Adductor Muscle of *Venus mercenaria* Stained with Phosphotungstic Acid. (Courtesy C. E. Hall, M. A. Jakus, and F. O. Schmitt, Massachusetts Institute of Technology.)

The fibrils are ribbon-like and usually dry during preparation with a flat side against the supporting film. Stained spots in consecutive cross bands are aligned at definite angles to the fiber axis, as indicated by the angles α and β, drawn on the electron micrograph. These angles satisfy

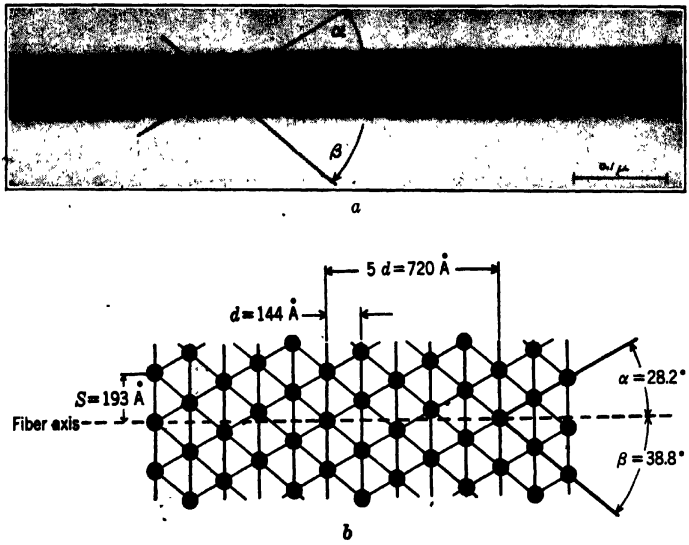


FIG. 9·21. (*a*) Fibril from Adductor Muscle of *Venus mercenaria*, Showing Two-Dimensional Lattice. (*b*) Drawing of Stained Lattice Structure from Electron Microscope Data. (Courtesy C. E. Hall, M. A. Jakus, and F. O. Schmitt, Massachusetts Institute of Technology.)

the relation $\tan \alpha = \frac{2}{3} \tan \beta$. The disposition of spots in the cross bands and diagonals forms a remarkably perfect geometrical pattern, which is shown diagrammatically in Fig. 9·21*b*. The pattern is laterally asymmetrical. If through any spot a line is drawn parallel to the fiber axis, it passes through a series of spots spaced five bands apart. The length of the unit cell is therefore 5·144 = 720 A.U., as indicated in the figure. This has been verified by x-ray diffraction observations. On the other hand, the x-ray observations do not yield a lateral period of 193 A.U. This suggests that not all the spots which are here observed lie in the same plane.

Figure 9·22 represents an electron micrograph of a cut section of guinea pig liver prepared in a manner commonly employed in light microscopy. The specimen was fixed with osmic acid and impregnated

with paraffin. This section, which is 0.2 micron in thickness, was obtained by E. Fullam[28b] of the Interchemical Corporation Laboratories with the aid of a high-speed knife.[28c]



FIG. 9·22. Section of Guinea Pig Liver Fixed with Osmic Acid and Impregnated with Paraffin. (Courtesy E. Fullam, Interchemical Corporation Laboratories.)

A special role in microscopy has long been assumed by the diatoms. The silica skeletons of these common pond algae have long served as test objects for the light microscope; a large part of their highly regular structure lies just at the limit of resolution of this instrument. Figure 9·23 illustrates the variety and beauty of the forms observed. A particular diatom frequently used as test object, *Pleurosigma angulatum*, has been the subject of two recent investigations utilizing the electron microscope.[29] The structures arrived at in the two investigations, both utilising stereomicrography, differ. In any case, they suggest an unsuspected degree of complexity.

As a last point in connection with the biological applications of the electron microscope the question may be raised whether living substance can be imaged with the electron microscope. This question has recently been answered in the affirmative by von Ardenne.[30] The technique of recording the micrograph, in such a manner as to minimize the

[28b] See Fullam and Geesler, reference 31*b*.
[28c] See O'Brien and McKinley, reference 31*c*.
[29] See Müller and Pasewaldt, reference 32, and Hamly and Watson, reference 33.
[30] See reference 34.

exposure of the living specimen to the electron beam, has already been described at the end of section 8·3. The spores of *Bacillus vulgatus*, the potato bacillus, were chosen as specimens because of their known extraordinary resistance to heat and drought.[31] These spores were



FIG. 9·23. Diatoms. (Specimen by Courtesy of Mr. Wardlaw M. Hammond, Philadelphia.)

placed on aluminum oxide films, which are mechanically stronger than the more usual collodion films. After exposure, the specimen support was removed from the microscope and immersed in nutrient. When a picture of the same field of view was taken later — considering a particular example — most of the spores were found to be still in their original position. A considerable number were missing, however, in the region of the field which had been protected from the electron beam during focusing before the first exposure, whereas almost none were missing in the unprotected region. It is reasonable to assume that these, or at least most of them, germinated in the nutrient and floated off. This interpretation was reinforced by the fact that at least two of the spores in the protected region were shown germinating in the second

[31] See v. Ardenne and Friedrich-Freksa, reference 35.

micrograph. These pictures thus constitute positive proof that *at least a large fraction of the spores in the protected region was still living after the first micrograph had been taken.*

The above work was done with an operating voltage of 180 kilovolts and a magnification of 500. The high voltage reduces the energy absorption in the specimen and increases the lethal dosage;[32] the low magnification diminishes the amount of charge per unit area that has to pass through the specimen to leave a satisfactory record on the photographic plate. When the magnification was increased to 1000 and the exposure increased correspondingly, all the spores in the second micrograph were in the same position as in the first, nor was there any sign of germination. Evidently the exposure even in the protected region had been enough to kill the spores. It is thus clear that, although living substance can survive being imaged in the electron microscope, it requires both a highly resistant specimen and very favorable conditions of operation.

9·3. Chemical Research. In the chemical field, as has been mentioned earlier in this chapter, the electron microscope bridges the gap between the lattice separations in crystals, which are a few angstrom units in length and may be measured with high accuracy by the methods of x-ray and electron diffraction, and larger particles and structures, of the order of a micron in diameter, that are readily observed with the light microscope. A vast number of problems of great scientific and industrial import are concerned with objects in the intermediate range.

Observations on large organic molecules belong close to the bottom of this range. Micrographs have been obtained of individual molecules of hemocyanin, hemoglobin, edestin, glycogen, and several other proteins.[33] Their molecular weights range from about 7,000,000 to 300,000; their diameters, correspondingly, from well over 200 A.U. to less than 100 A.U. if the molecules are assumed to be spherical and to have a density of 1.5, corresponding to that of the compact material. The diameter of the smallest organic molecule which can be satisfactorily observed in this manner is not determined by the resolving power of the instrument, but by the least contrast required to distinguish the image of the molecule from its background. Thus von Ardenne[33] estimates that the smallest observable organic molecule should have a diameter of about 40 A.U. and a molecular weight of 40,000. The direct imaging of large organic molecules can be of value in several

[32] This may be defined as the amount of charge that must be incident on unit area of the specimen to kill half of the living cells.

[33] See von Ardenne, reference 36, Stanley and Anderson, reference 37, and Husemann and Ruska, reference 38.

respects: It provides a check on the value of the size and molecular weight determined by indirect methods; it indicates the shape of the molecule; and it may throw light on the mode of aggregation of the molecules in a continuous film. Figure 9·24 shows hemocyanin, the blue coloring matter found in the blood of invertebrates, applied in a low concentration to the supporting film. Here the individual molecules are readily distinguished; at higher concentrations they form a continuous film superposed on the collodion film.

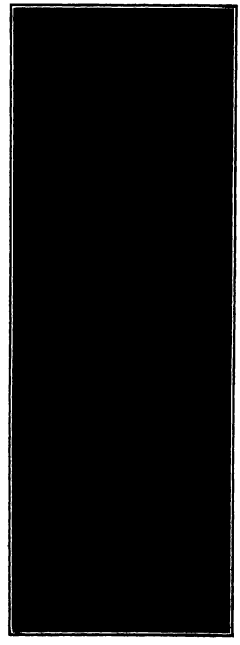


FIG. 9·24. Hemocyanin from *Limulus polyphemus*. (Courtesy *J. Biol. Chem.*, reference 37.)

Colloidal suspensions, smokes, dusts, pigments, clays, and insecticides are but a few of the materials which fall into the range of application of the electron microscope. The information to be gained with the electron microscope relates primarily to the size distribution and the shapes of the individual particles — the factors which determine in large measure the chemical activity and physical behavior of the materials in their various applications. It is well to note, in this connection, that, for the true shape of the particle to be recognized, the resolution of the instrument must be much finer than the diameter of the particle. Suppose that the particle is a regular polygon. Then, in the image, each corner will appear rounded off within an area equal to the circle of confusion drawn about the true corner point. If it is assumed, with von Borries and Kausche,[34] that the portion of each polygon side which still appears straight in the image must be at least equal in length to the chord of the curved contour of the adjoining corners in order that the particle be recognized as being a polygon of the right number of sides, the ratio of the particle diameter d' (defined as the diameter of a circle of equal area) and of the instrument resolution δ is shown, for polygons with different numbers of sides, in Fig. 9·25. Roughly, the particle diameter must exceed the resolution by a factor equal to the number of its sides. The meaning of the criterion employed is shown graphically in Fig. 9·26, indicating the image contours of various polygonal

[34] See reference 39.

particles whose shape would be just recognizable with the same resolution δ.

A careful determination of particle shape and size distributions in a number of gold sols was made by von Borries and Kausche. They

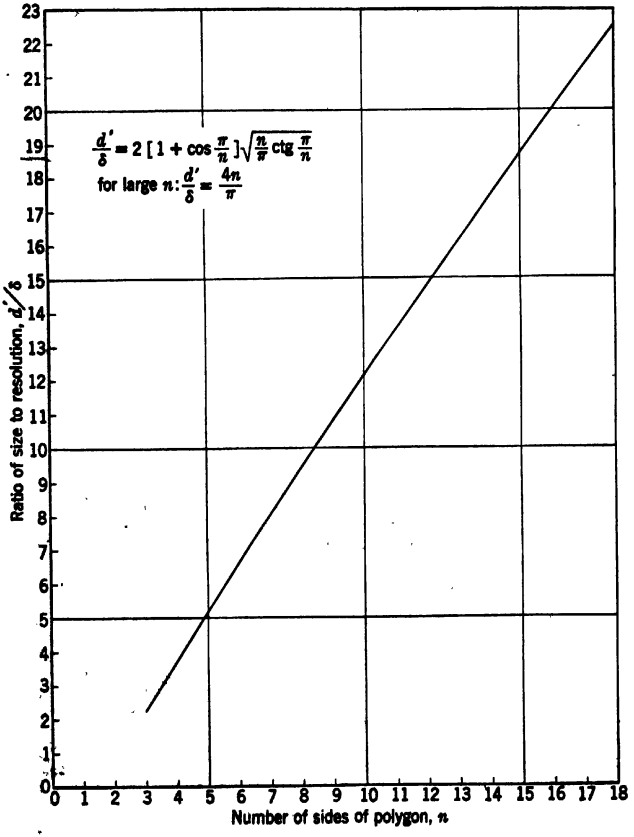


FIG. 9·25. Ratio of Least Diameter of Polygon Permitting Recognition of Shape to Limit of Resolution as Function of the Number of Sides of the Polygon. (Reference 39.)

found the particles to be octahedral. The distributions were Gaussian, with maxima for particle diameters ranging from 15 to 50 millimicrons.

Of immediate practical importance have been the studies of carbon blacks, carried out, in particular, by the Columbian Carbon Research

Laboratories.[35] These carbon blacks are employed as fillers to increase the mechanical strength of both natural and synthetic rubbers. Figure 9·27 shows the appearance of a typical carbon dispersion. The individual particles are spheroidal, with smooth edges. Size distributions measured with the electron microscope indicated, in the case of the finer carbons, a much smaller mean particle diameter than had been inferred from particle counts on measured quantities of material with the aid of the light ultramicroscope, the principal errors of the latter method all operating to yield too large mean diameter values. Thus the mean diameter of the carbon whose distribution is represented in Fig. 9·28 turned out to lie between 50 and 60 millimicrons when determined by count. The importance of knowing the size distribution of the carbon black employed in rubber reinforcement is well illustrated by the two curves in Fig. 9·29,[36] which show the wearing quality (as measured by the product of rupture energy and shore hardness) and the percentage rebound of the reinforced rubber as function of the surface area per unit mass of the carbon or, roughly, the inverse of the mean diameter of the carbon particles. The rebound is one of the factors determining the heavy-duty performance of the material, to be considered in particular with truck and bus tires. It is seen that the wearing quality increases uniformly, though at a decreasing rate, with the fineness of dispersion of the carbon black, whereas the rebound shows a steady decrease. The selection of the best material is determined by the relative importance of the two factors for the purpose in question. A further study of the coordination of the carbon particle size distributions and the physical properties of vulcanized rubber was carried out by Schoon and Koch.[37] These authors regarded the particles as hexagonal or circular platelets rather than as spheroids. More recent

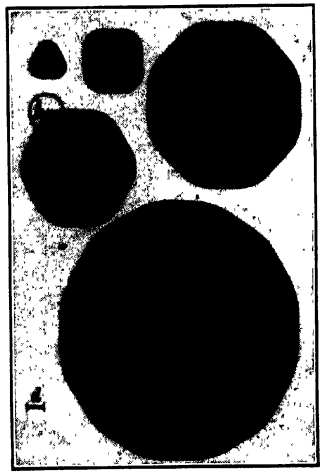


FIG. 9·26. Image Contours of Various Polygons of Barely Recognizable Shape for Identical Instrumental Resolving Power. (Reference 39.)

[35] See Columbian Carbon Company, reference 40.

[36] See Wiegand and Ladd, reference 41.

[37] See reference 42.

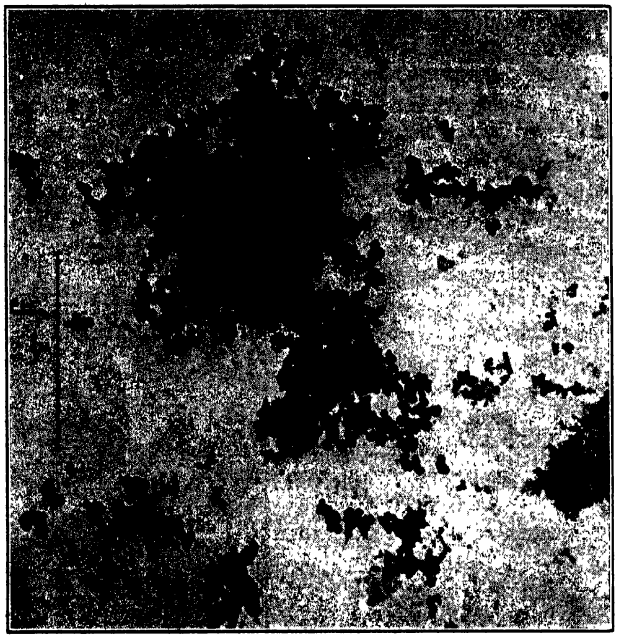


FIG. 9-27. Carbon Black from a Natural Gas Flame.

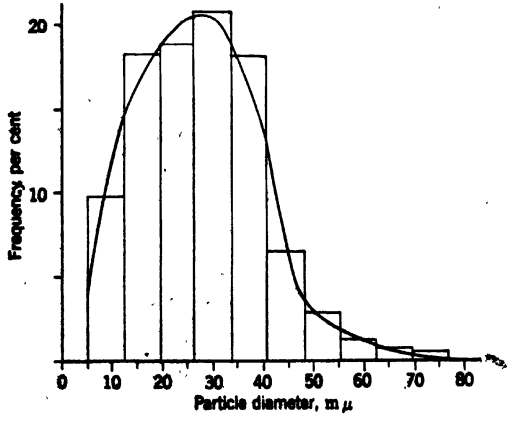


FIG. 9-28. Frequency Distribution of the Diameters of Carbon Particles in "Micronex" Carbon Black. (Courtesy Columbian Carbon Company, reference 40.)

work on some of the coarser carbon blacks has confirmed the spheroidal character of the particles.

Other smokes which have attracted considerable attention[38] are the metallic smokes, which constitute health hazards in some industries.

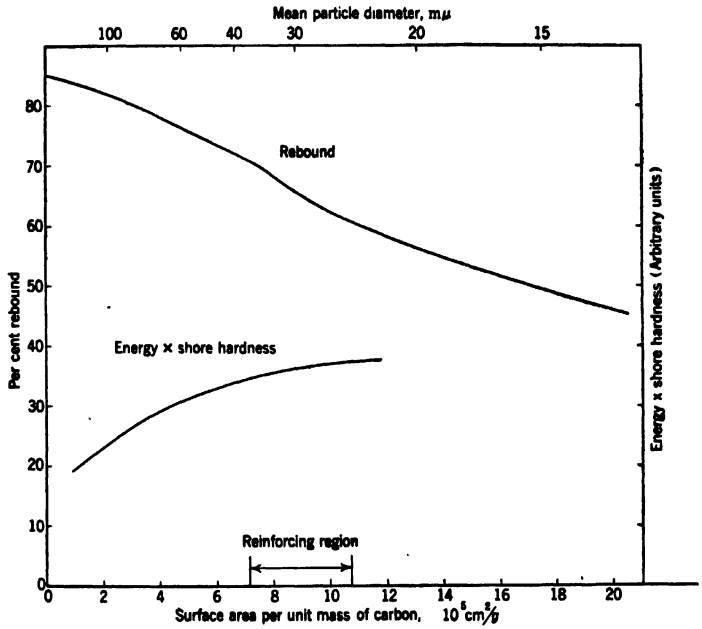


FIG. 9·29. Wearing Quality (Rupture Energy × Shore Hardness) and Rebound of Rubber as Function of the Surface Area per Unit Mass of the Carbon Filler. (Reference 41.)

A few of these, with their characteristic, highly distinctive particle shapes, are shown in Figs. 9·30–9·32. The magnesium and zinc smokes were obtained by burning the metal in air. They consist of oxides of the respective metals. The aluminum smoke was derived from an electric arc. The spherical particles are presumably aluminum particles with an oxide coating. The zinc oxide employed as a pigment[39] presents an appearance somewhat different from that of the smoke (Fig. 9·33). The impossibility of distinguishing the shapes or estimating the sizes of such particles is well illustrated by the ultraviolet micrograph

[38] See, for example, Mahl, reference 43, and Barnes and Burton, reference 44.
[39] Kadox Black Label of the New Jersey Zinc Company of Pennsylvania.

of the same material prepared by the New Jersey Zinc Company (Fig. 9·34). The particle size is here of great importance, since it determines, very largely, the covering power of the pigment. The covering power of a pigment increases with decreasing particle size until a point in the

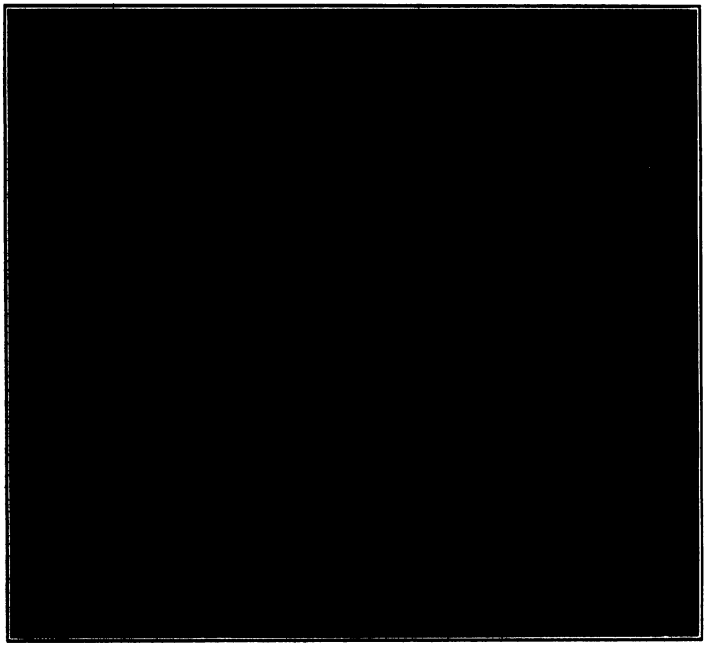


FIG. 9·30. Magnesium Smoke (MgO).

neighborhood of the limit of resolution of the light microscope is reached. Beyond this point it decreases again. Schmieder,[40] with the aid of the electron microscope, has studied the coordination between particle size and covering power for certain azo pigments in the range of very small particle sizes, which is utilized, in particular, in multicolor printing.

A large surface area, or fine subdivision, of the active material is desirable in many insecticides (Fig. 9·35), disinfectants (Fig. 9·36), and catalytic reagents (Fig. 9·37). The thin-plate form of a contact poison such as lead arsenate is also favorable from the point of view of facilitating adhesion to foliage. Both in mercurochrome, represented in Fig. 9·36, and in the sample of activated catalyst alumina shown

[40] See reference 45.



FIG. 9·31. Zinc Smoke (ZnO).

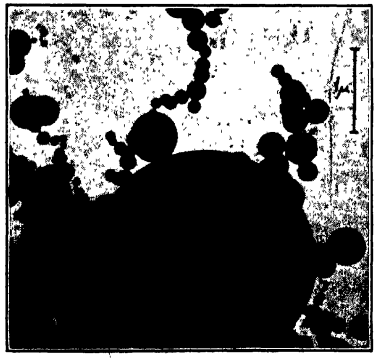


FIG. 9·32. Aluminum Smoke (Al + Al_2O_3). (Courtesy Stamford Research Laboratory, American Cyanamid Company. Reference 44.)

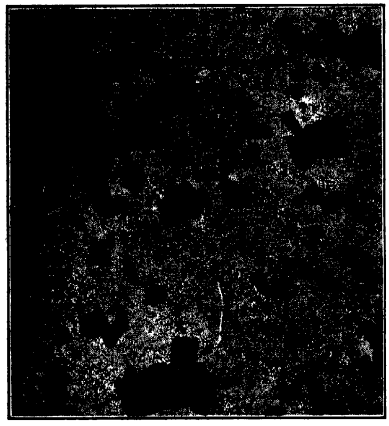


FIG. 9·33. Zinc Oxide Pigment (Kadox). (Courtesy New Jersey Zinc Company of Pennsylvania.)

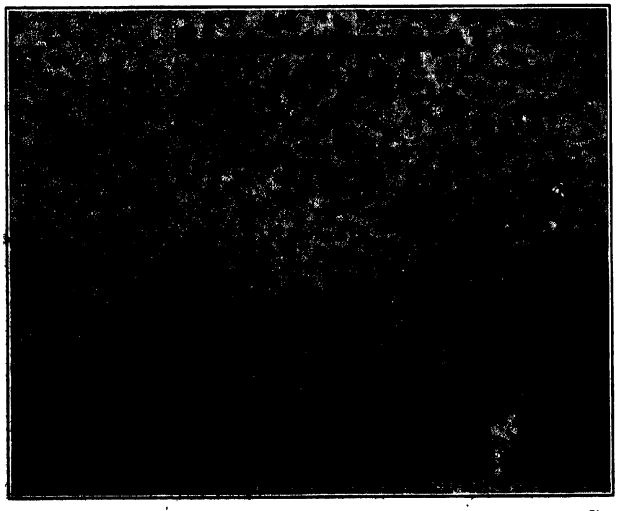


FIG. 9·34. Ultraviolet Micrograph of Pigment Shown in Fig. 9·33. (Courtesy New Jersey Zinc Company of Pennsylvania.)

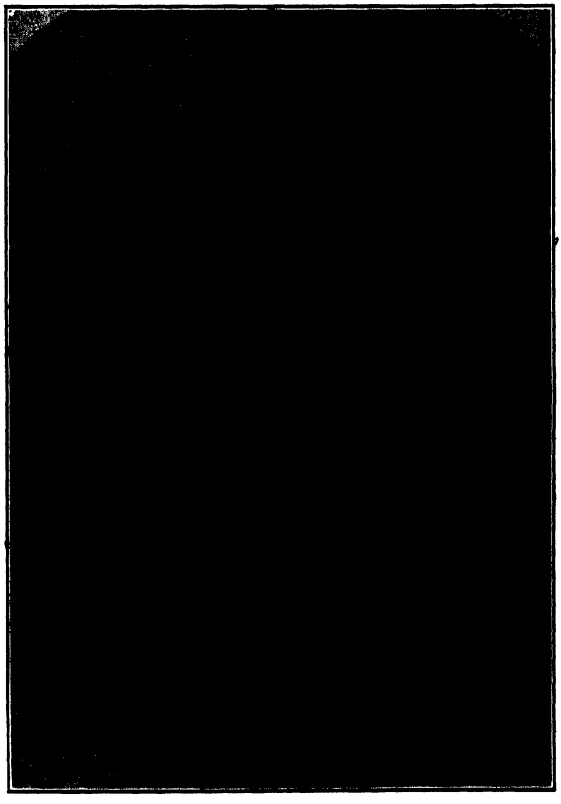


FIG. 9-35. Lead Arsenate, an Insecticide.

in Fig. 9·37 the active substance is dispersed in particles of extraordinarily small size. Numerous catalysts have been examined with the electron microscope.[41]

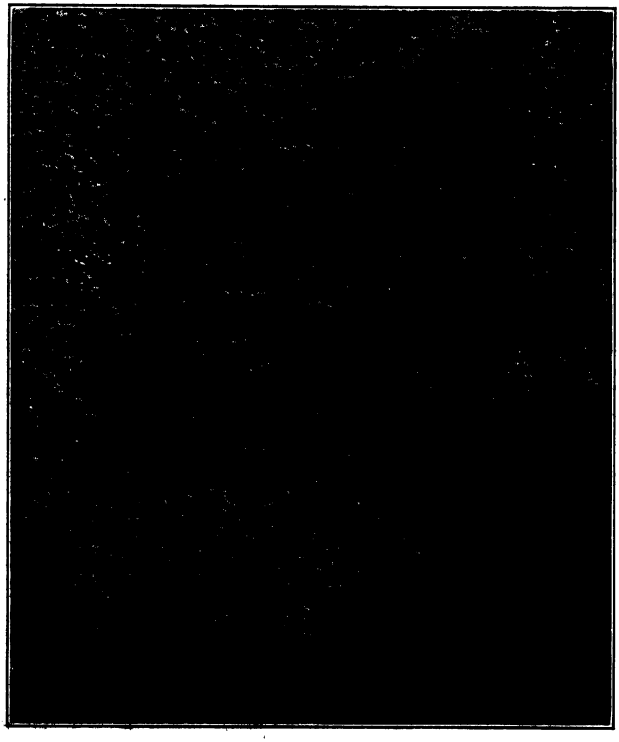


FIG. 9·36. Mercurochrome.

Clay minerals have been studied in considerable detail with the electron microscope. In particular, the characteristic appearance of the elementary particles of numerous clay minerals has been demonstrated and described by Shaw and his coworkers.[42] The structural properties of montmorillonite, a clay employed in the bleaching of oils because of its high adsorptive capacity, have been investigated by

[41] See von Ardenne, reference 13, pp. 324–332.

[42] See Shaw and Humbert, reference 46, and Marshall, Humbert, Shaw, and Caldwell, reference 47.

electron microscopy[43] and electron diffraction[44] in the same instrument, comparing them with those of caolinite. Eitel and Schusterius[45] made determinations of the size distributions in fractions of caolinite minerals. The stereophotogrammetric method of determining thicknesses, described in section 8·6, was employed to obtain estimates of the thickness

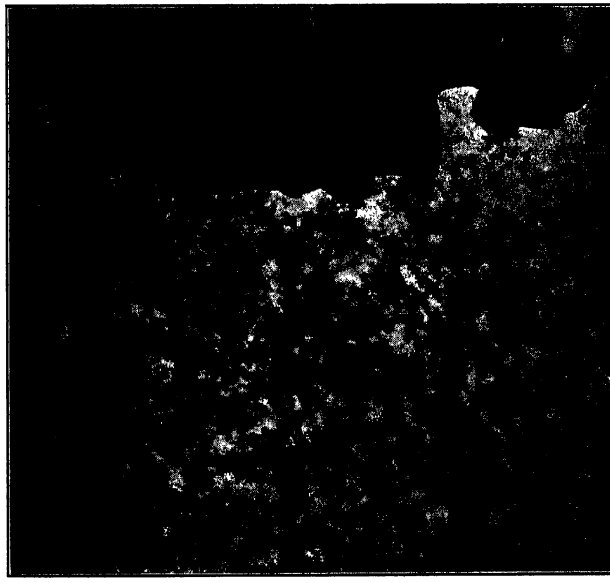


FIG. 9·37. Activated Alumina.

of caolin flakes.[46] Characteristic hexagonal platelets of caolinite are shown in Fig. 9·38. Finally, Eitel, Müller, and Radczewski followed out, step by step, the disintegration and eventual recrystallization of natural and synthetic caolin as well as of pholerite in the baking process.[47]

The authors last mentioned have also made a study of the hydration processes which result in the setting of cement, that is, the hydration of quick lime,[48] tricalcium silicate,[49] and tricalcium aluminate.[50] Tri-

[43] See Eitel and Radczewski, reference 48, and Humbert and Shaw, reference 49.
[44] See O'Daniel and Radczewski, reference 50.
[45] See reference 51.
[46] See Eitel and Gotthardt, reference 52.
[47] See Eitel, Müller, and Radczewski, reference 53.
[48] See Radczewski, Müller, and Eitel, reference 54.
[49] See Radczewski, Müller, and Eitel, reference 55.
[50] See Radczewski, Müller, and Eitel, reference 56.

calcium silicate is found to decompose, in the presence of water, into hemispherical particles of calcium hydroxide and needle-shaped crystallites of a calcium hydrosilicate of indefinite composition. The crystallites appear to form groups which, eventually, develop into larger, regular crystals.

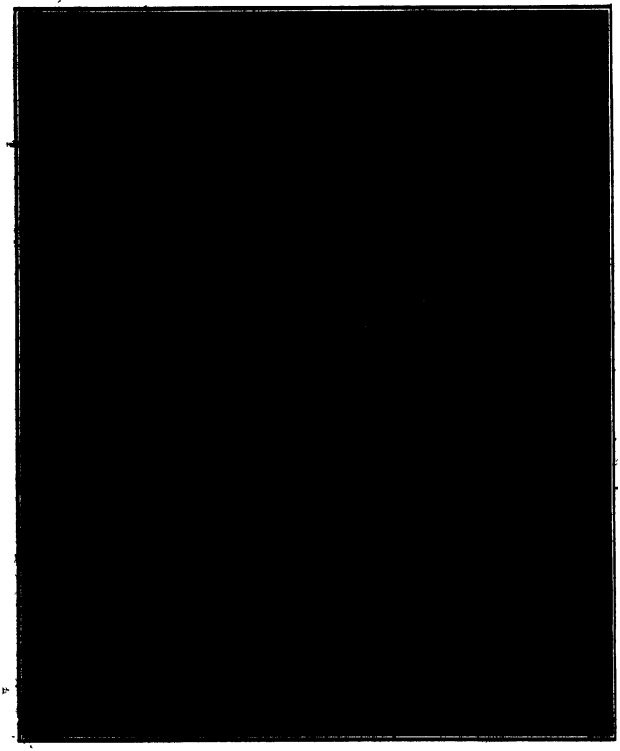


FIG. 9-38. Funkhouser Clay.

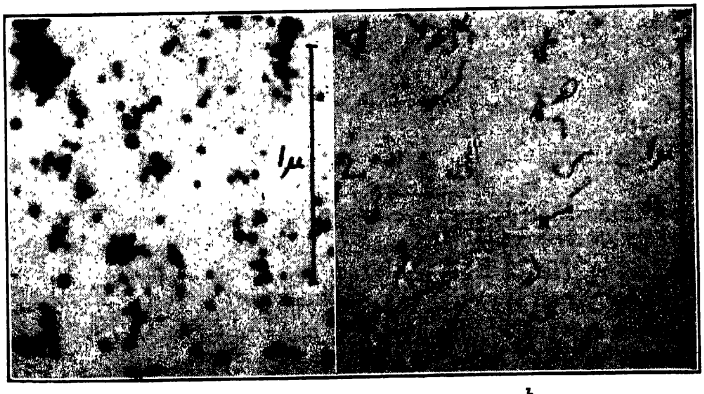
Photographic phenomena, also, have been investigated with the aid of the electron microscope. Hall and Schoen[51] have studied the reduction of silver bromide crystals by light and by the electron beam itself as well as the development of exposed crystals by various developers. It was shown that normal development causes thin silver fibers to grow out of the bromide crystals. Bundles of these are left when the bromide crystal is removed by fixing and, thus, constitute the developed image.

[51] See reference 57.



a b

FIG. 9-39. Silver Bromide Crystal Exposed to Electron Beam (*a*) before Development; (*b*) after Development and Fixation. (Courtesy Research Laboratories of the Eastman Kodak Company, Rochester, N. Y. Reference 57.)



a b

FIG. 9-40. Lippman Crystals (*a*) before Development; (*b*) after Development and Fixation. (Courtesy Research Laboratories of the Eastman Kodak Company, Rochester, N. Y. Reference 57.)

Figure 9·39 shows the appearance of a typical silver bromide crystal and of the silver residue after development and fixation. In the case of the very small Lippman crystals — of the order of 300 A.U. in diameter — each crystal forms a single silver ribbon which is as much as five times as long as the original crystal (Fig. 9·40).

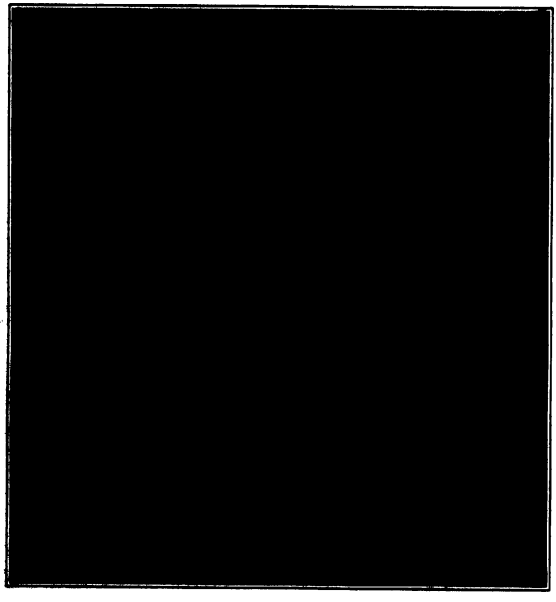


FIG. 9·41. Sodium Laureate Curd. (Reference 58.)

Numerous materials may be spread in the form of continuous films or fine networks over the object support. Examples of these are the soaps and plastics. In the study of sodium laureate curds Marton, McBain, and Vold[52] found the material to be distributed in a tight network of branching fibers containing quantities of minute granules (Fig. 9·41). The resulting large number of small capillary spaces in the curd facilitates the retention of water even at low humidity.

Among the plastics, polystyrene (Fig. 9·42), important in the electrical industry, and Koroseal or polymerized vinyl chloride (Fig. 9·43), a synthetic rubber, may serve as examples. Such pictures contain a large amount of information that awaits exact interpretation. Many of the dark specks in the Koroseal fiber are only about 30 A.U. in diameter,

[52] See reference 58.

corresponding in magnitude to a moderate-sized protein molecule! It is only reasonable to expect that this method of investigation will play a considerable role in obtaining an understanding of the physical properties and the structure of these materials. Natural rubber has, of

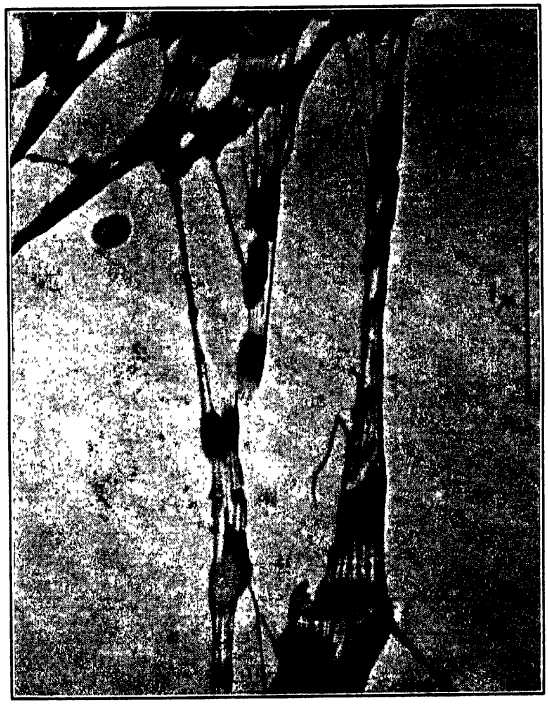


FIG. 9-42. Polystyrene.

course, also been the subject of electron-microscopic study.[53] A micrograph of natural latex is shown in Fig. 9-44.

Both the technique of sectioning[54] and that of disintegration with the aid of chemical reagents[55] or mechanical devices are adopted for investigating the structure of cotton, silk, and rayon fibers with the aid of the electron microscope. The first method yields information regarding the normal arrangement of the fiber; thus cotton and rayon hairs have

[53] See, for example, v. Ardenne and Beischer, reference 59.
[54] See Ruska, reference 60.
[55] See Ruska and Kretschmer, reference 61.

been found to be made up of a dense bark surrounding a spongy marrow containing large numbers of intercommunicating, spherical cavities. The method of disintegration, the second method, reduces the fibers to their structural components. These *micelles*, in the case of cotton,

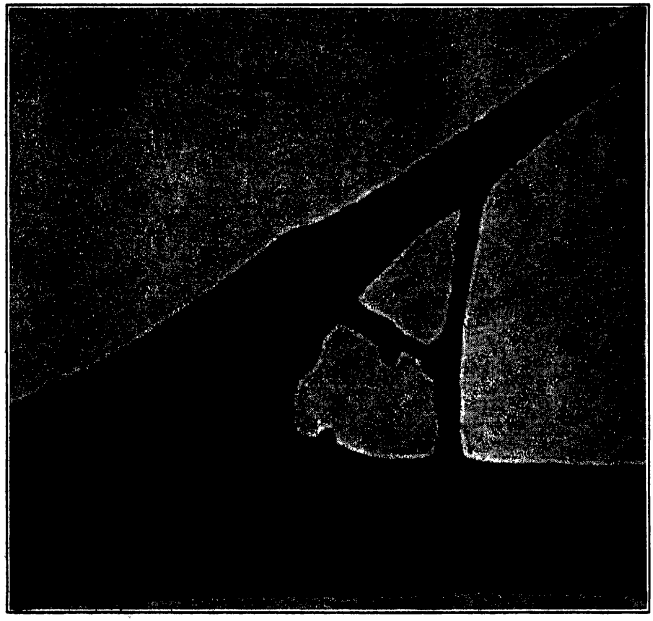


FIG. 9·43. Koroseal (Polymerized Vinyl Chloride). (Specimen by Courtesy of Goodrich Rubber Company, Akron, O.)

appear to be long, thin threads about 50 A.U. in diameter. Similar fibrils have been observed in the cementing material extracted from wood pulp.[56] Considerable caution is required in interpreting the observations on relatively thick fibers, since, here, the electron beam readily causes a swelling and eventual charring of the fiber.[57] This may account for some of the earlier observations on fiber sections. Figure 9·45 may be regarded as representative of the results obtained more recently in the field of textile fibers.[58]

The electron microscope is particularly well suited to the study of

[56] See Sears and Kregel, reference 62.
[57] See Mahl, reference 63, and Barnes and Burton, reference 64.
[58] See Barnes and Burton, reference 65.

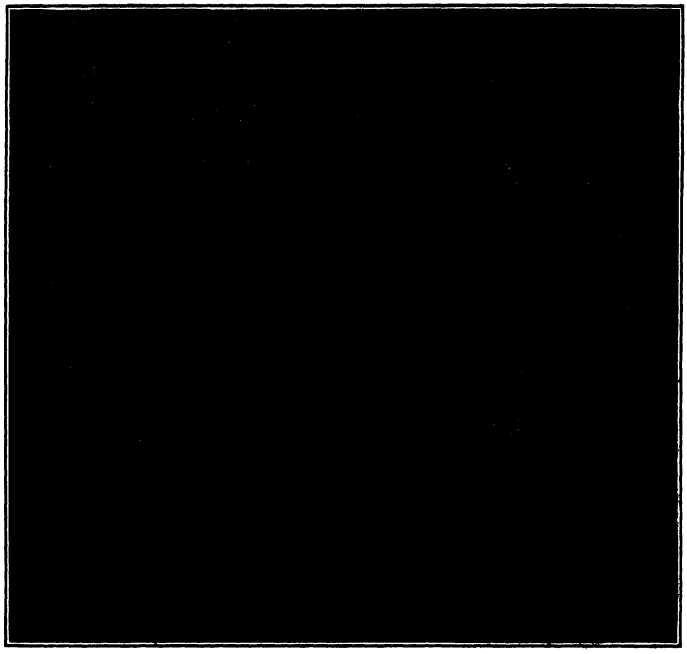


FIG. 9-44. Natural Rubber Latex.



FIG. 9-45. Cotton Fibers Disintegrated in Water. (Courtesy Stamford Research Laboratories, American Cyanamid Company. Reference 65.)

evaporated films of metals, which may be deposited directly on the collodion film acting as specimen support. Earlier publications — excepting a study of evaporated antimony films[59] — confined themselves primarily to the reproduction of a number of isolated pictures,[60]

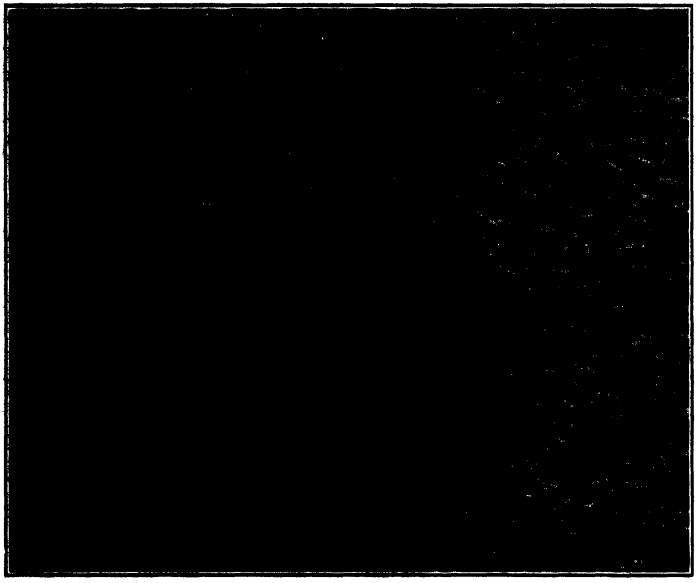


FIG. 9·46. Evaporated Silver Film.

eventually accompanied by electron diffraction patterns.[61] However, Picard and Duffendack[62] have published recently a systematic examination of a large number of different metal films, varying the thickness. In all cases the films are found to consist of agglomerates separated by clear interstices. The initial breaking up of a molecular layer into droplets may be ascribed to the great surface tension of the small droplets formed. The metals may be subdivided, however, into two classes, on the basis of their surface mobility, the adhesion of the metal atoms to the metal agglomerates being in all cases stronger than to the collodion substrate. Only those with moderate surface mobility form good continuous films, showing initially regularly distributed agglom-

[59] See Ruedy, reference 66.
[60] See Mahl, reference 43.
[61] See Hillier, Baker, and Zworykin, reference 67.
[62] See reference 68.

erates separated by channels of approximately constant width. Examples of such metals are aluminum, antimony, copper, gold, magnesium, and silver.

Figure 9·46 shows a typical micrograph of a silver film. As the films become thicker the agglomerates increase in size and thickness. Even-

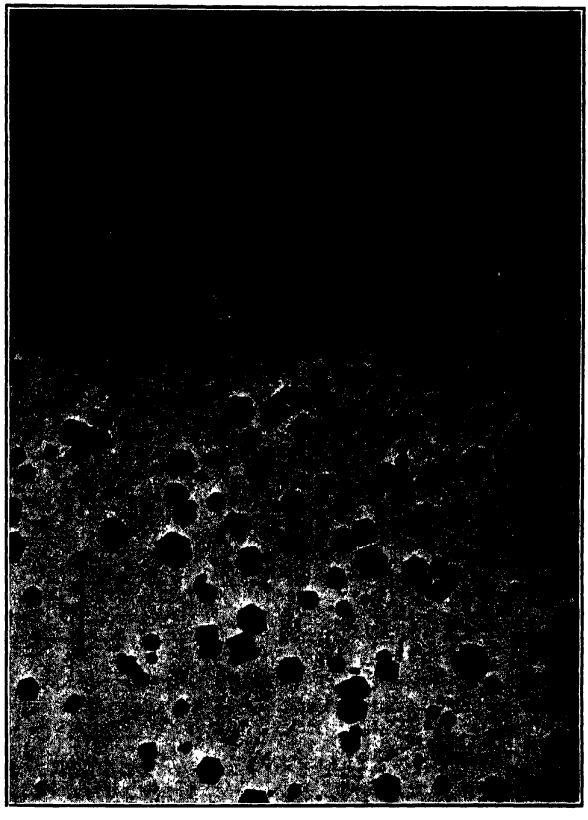


FIG. 9·47. Cadmium Evaporated on a Collodion Film.

tually the clear channels disappear and the substratum is covered completely. With zinc and cadmium, however, the case is different. With these metals the surface mobility on collodion (and on glass) is so great that continuous films are not formed at any time. The evaporated metal collects in scattered crystals varying in size. If, instead of a

plain collodion film, an evaporated aluminum or copper film is employed as substratum, continuous films of zinc or cadmium may eventually be formed. Even here, however, the free areas between the crystalline agglomerates are very large compared to those obtained with the first class of metals. Figure 9·47 shows the regular hexagonal forms obtained when cadmium is evaporated on collodion.

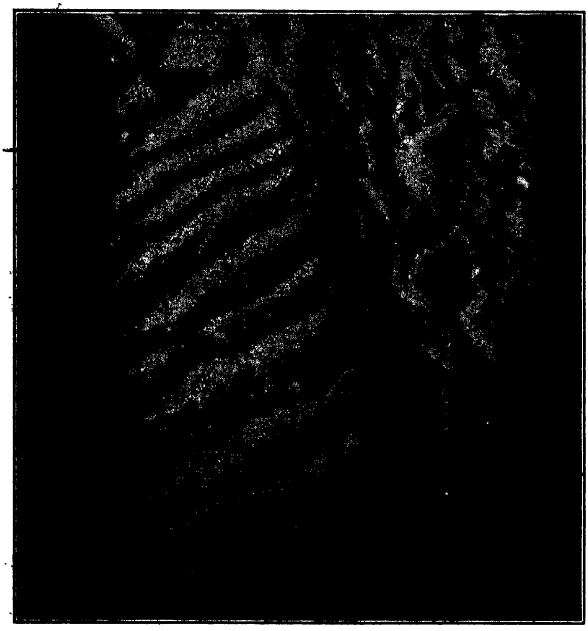


FIG. 9·48. Deformed Calcite Crystal, Showing Slip Bands. (Courtesy Physical Research Laboratory, Dow Chemical Co., Midland, Mich. Reference 70.)

In some cases the observation of the contours of objects much too dense to transmit electrons proves of value. Thus Haefer[63] employed the electron microscope to measure the radii of curvature of extremely fine etched tungsten points and utilized the data so obtained for an experimental check of the wave-mechanical theory of field emission. The replica techniques described in section 8·1 also prove valuable in the study of the surfaces of bulky specimens. Thus Fig. 9·48 shows the surface structure of a deformed calcite crystal, the deformation being

[63] See reference 69.

a consequence of the pressure applied in forming the mold preliminary to preparing the replica.[64] A replica micrograph of surface patterns of glazed porcelain, obtained by another method which does not involve the use of high pressures,[65] is shown in Fig. 9·49. The dark masses in two of the corners represent silver residues due to inadequate washing of the replica. Recently the replica technique has also been applied to a study of the fracture process in glass.[66]



FIG. 9·49. Surface Pattern on Glazed Porcelain. (Courtesy *J. Applied Phys.*, reference 71.)

9·4. Metallurgy. The replica technique last mentioned, and described in detail in section 8·1, was developed primarily to adapt the electron microscope to the study of metallographic specimens. It was first applied to problems related to the structure and the hardenability of steels by Mehl, who made a careful study, in particular, of pearlite.[67] This is a eutectoid form of steel which tends to crystallize out as steel

[64] See Heidenreich and Peck, reference 70.
[65] See Zworykin and Ramberg, reference 71.
[66] See Gölz, reference 72.
[67] See Mehl, reference 73.

is cooled through the range between 660 and 580° C. Its formation is prevented by rapid quenching, resulting in hardened steel. Pearlite consists of lamellae of iron carbide or cementite (Fe_3C) immersed in pure iron or ferrite, the ratio of the total amount of cementite and ferrite remaining fixed. Since the cementite is dissolved less readily than iron by etching reagents, it stands out in thin ridges on etched steel surfaces.



FIG. 9·50. Fine Pearlite, Formed at 640° C. (Courtesy *J. Applied Phys.* References 71 and 73.)

Figure 9·50 is a micrograph of characteristic pearlite structure. The separations between lamellae here shown lie close to the limit of resolution of the metallographic light microscope. As the temperature of formation is reduced, the lamellar spacing becomes smaller. Mehl was able to demonstrate that, contrary to some earlier opinions, the characteristic pearlite structure was maintained even when the lamellar spacing was reduced to 300 A.U. Under certain circumstances the cementite lamellae tend to break up, forming isolated nodules of cementite. Such a conversion of the lamellae into a nodular, or "spher-



FIG. 9.51. Pearlite — from Negative Formvar Replica. (Courtesy Research Laboratories of the General Electric Co., Schenectady, N.Y.)



FIG. 9·52. Lamellar Pearlite in 0.98 Per Cent Carbon Hot Roll Steel. (Courtesy Physical Research Lab., Dow Chemical Co., Midland, Mich. Reference 70.)

oidized" form, is obtained by prolonged tempering of pearlite steels, resulting in a form which is particularly favorable for machining. In these and other modifications of steel the iron carbide structure has been examined with the electron microscope, revealing detail not previously observed. Additional pictures of pearlite, obtained by two other replica methods, are shown in Figs. 9·51 and 9·52.

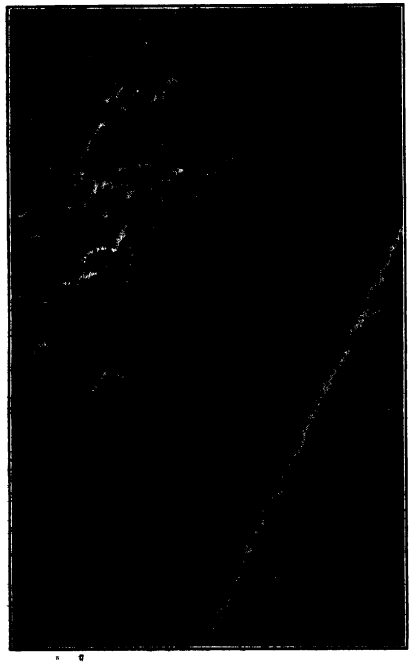


FIG. 9·53. Twinning and Precipitation in Copper-Beryllium Alloy. Specimen from D. Harker, General Electric Co. (Courtesy Physical Research Lab., Dow Chemical Co., Midland, Mich. Reference 70.)

Numerous nonferrous alloys have been studied, by Heidenreich and his associates,[68] in particular employing the polystyrene-silica method of forming replicas.[69] Figures 9·53 and 9·54 are examples of the fine micrographs thus obtained. The first of these shows twinning in a copper-beryllium alloy; the fine specks represent, presumably, pre-

[68] See Heidenreich and Peck, reference 70.

[69] See section 8·1.

cipitated beryllium. The other is a micrograph of a magnesium-aluminum alloy. The taking of stereopictures was found helpful in interpreting the replica micrographs.[70] Particles separated by as little as 75 A.U. could be distinguished on the latter.

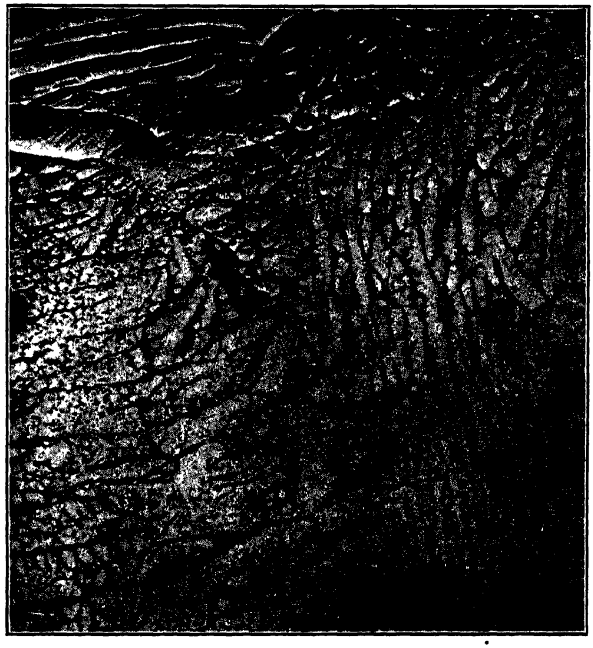


FIG. 9·54. Magnesium-Aluminum Alloy. (Courtesy Physical Research Laboratory, Dow Chemical Co., Midland, Mich.)

Aluminum and aluminum alloys have also been studied with great success by Mahl,[71] employing oxide-film replicas.[72] These constitute the earliest effective replica technique. A typical micrograph of an etched aluminum surface, obtained in this manner, is shown in Fig. 9·55. More recently, Mahl and Stranski[73] have analyzed the etching process for aluminum with the aid of oxide-film electron micrographs, demonstrating the essential role played by the formation of a "two-dimen-

[70] See Heidenreich, reference 74.
[71] See references 75–77.
[72] See section 8·1.
[73] See reference 78.

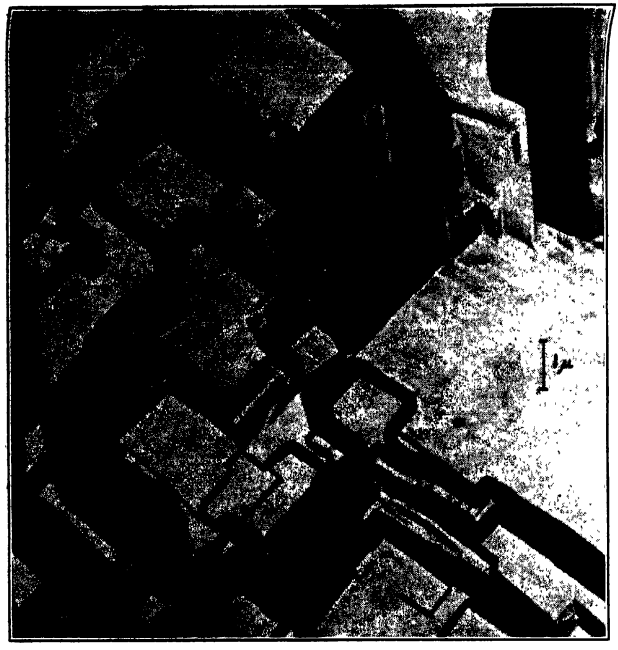


FIG. 9-55. Aluminum Etched with Nitric Acid; Oxide-Film Replica Micrograph. (Courtesy Physical Research Laboratory, Dow Chemical Co., Midland, Mich.)

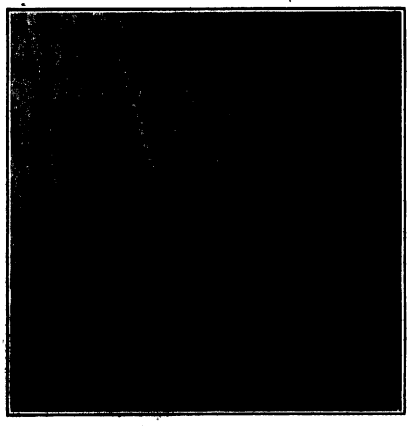


FIG. 9-56. Grinding Marks on a Steel Surface Finished on a Chamois Wheel with Levigated Alumina. (Courtesy Physical Research Laboratory, Dow Chemical Co., Midland, Mich. Reference 70.)

sional " film of aluminum oxide at the interface of the metal and the wet etching reagent.

The replica techniques are, of course, also applicable to the study of unetched surfaces, such as those of machined parts and bearings, and may here yield information which is withheld by the light microscope owing to its limited resolution. Thus Fig. 9·56 shows grinding marks on a steel surface. A study of hardened and unhardened steel surfaces after coarse grinding, fine grinding, and polishing has been made by von Borries and Janzen[74] observing the specimens, inclined at a glancing angle, directly with the aid of reflected electrons as described in section 8·7. Even though the micrographs so obtained are scarcely comparable to those recorded with the replica technique, the gain in resolution proved sufficient to throw new light on the action of the working processes on the surfaces.

9·5. Prospects. The survey of past accomplishments of the electron microscope which has been given is — at best — sketchy. Many important and valuable researches have not been touched, others barely mentioned. Even so it is clear that, within the very few years that the electron microscope has been available for research, it has left its imprint on most branches of experimental science. This process may be expected to proceed at an accelerated pace. The most striking impression gained in examining the literature of electron microscopy is the progressively increasing number of papers — in the face of external conditions unfavorable to publication — and, particularly, the improving quality of the research described. To an ever greater extent this research is becoming quantitative, rather than purely qualitative.

In the future, the pace may be expected to be accelerated by the introduction of simpler, more compact instruments,[75] on the one hand, and by improvements in specimen preparation technique on the other. The latter development has been most striking, possibly, in the field of metallurgy. However, in other fields also, such as the preparation and examination of biological sections and the study of chemical reactions within the microscope, progress has been made and further progress is to be expected. Another avenue of increasing the utility of the electron microscope is in combining its use with that of electron diffraction[76] and other analytical techniques which are in the course of development. Not the least factor in rapidly extending the usefulness of the new tool will be the quick interchange of new experience and information among active workers, facilitated by organizations such as the Electron Microscope Society of America, formed in Chicago in November, 1942.

[74] See reference 79.

[75] See, for example, sections 5·1 and 3·3.

[76] See section 8·8.

As far as the electron microscope itself is concerned, it may be stated with some confidence that no improvements in the limit of resolution by orders of magnitude are to be looked for. When such improvements occur, they are likely to be the consequence of a new departure, just as the electron microscope has represented a new departure relative to the light microscope. The principal immediate gains are more likely to lie in the field of greater ease of manipulation and a more uniformly high quality of the results.

No such statement can be made, however, with regard to the future results of electron-microscope research. He would indeed be daring who would attempt to circumscribe this work and its consequences. Yet one thing is certain. The electron microscope will have a material share in extending man's understanding and mastery over nature.

REFERENCES

1. L. MARTON, "Electron microscopy of biological objects," *Bull. acad. roy. Belg.*, Classe des Sciences, Vol. 20, p. 92, 1934; *Nature*, Vol. 133, p. 911, 1934.
2. L. MARTON, "Electron microscopy of biological objects," *Bull. acad. roy. Belg.*, Classe des Sciences, Vol. 23, p. 672, 1937.
3. W. M. STANLEY and T. F. ANDERSON, "A study of purified viruses with the electron microscope," *J. Biol. Chem.*, Vol. 139, pp. 325–338, 1941.
4. G. A. KAUSCHE, E. PFANKUCH, and H. RUSKA, "The making visible of plant virus in the supermicroscope," *Naturwissenschaften*, Vol. 27, pp. 292–299, May 1939.
5. E. PFANKUCH and G. A. KAUSCHE, "Isolation and supermicroscopic imaging of a bacteriophage," *Naturwissenschaften*, Vol. 28, p. 46, 1940.
6. S. E. LURIA and T. F. ANDERSON, "The identification and characterization of bacteriophages with the electron microscope," *Proc. Natl. Acad. Sci. U.S.*, Vol. 28, pp. 127–130, April 1942.
7. S. E. LURIA, M. DELBRÜCK, and T. F. ANDERSON, "Electron-microscope studies of bacterial viruses," *J. Bact.*, Vol. 46, pp. 57–76, July 1943.
8. R. H. GREEN, T. F. ANDERSON, and J. E. SMADEL, "Morphological structure of the virus of vaccinia," *J. Exptl. Med.*, Vol. 75, pp. 651–656, June 1942.
9. T. F. ANDERSON and W. M. STANLEY, "A study by means of the electron microscope of the reaction between tobacco mosaic virus and its antiserum," *J. Biol. Chem.*, Vol. 139, pp. 339–344, 1941.
10. G. A. KAUSCHE and H. RUSKA, "The detection of molecules of tobacco mosaic virus in the chloroplasts of virus-diseased plants," *Naturwissenschaften*, Vol. 28, p. 303, May 1940.
11. G. A. KAUSCHE, "On a substance inactivating the virus protein in the seed of *Nicotiana tabacum samsun*," *Biol. Zentr.*, Vol. 60, pp. 423–438, July–August 1940.
12. M. v. ARDENNE and G. PYLA, "Attempts at imaging foot-and-mouth disease virus with the Universal Electron Microscope," *Naturwissenschaften*, Vol. 28, pp. 531–532, August 1940.
13. M. v. ARDENNE, *Übermikroskopie*, J. Springer, Berlin, 1940.

14. D. G. Sharp, A. R. Taylor, D. Beard, and J. W. Beard, "Electron micrography of the western strain equine encephalomyelitis virus," *Proc. Soc. Exptl. Biol. Med.*, Vol. 51, pp. 206–207, 1942.
15. A. R. Taylor, D. G. Sharp, D. Beard, and J. W. Beard, "Electron micrography of the eastern strain equine encephalomyelitis virus," *Proc. Soc. Exptl. Biol. Med.*, Vol. 51, pp. 332–334, 1942.
16. S. Mudd and D. B. Lackman, "Bacterial morphology as shown by the electron microscope," *J. Bact.*, Vol. 41, pp. 415–420, 1941.
17. S. Mudd, K. Polevitzky, T. F. Anderson, and L. A. Chambers, "Bacterial morphology as shown by the electron microscope II," *J. Bact.*, Vol. 42, pp. 251–264, August 1941.
18. A. Jakob and H. Mahl, "Representation of structure in bacteria, in particular of capsules in anaerobic strains, with the electrostatic electron supermicroscope," *Arch. exptl. Zellforsch. Gewebezücht.*, Vol. 29, pp. 97–104, 1940.
19. A. Lembke and H. Ruska, "Comparative microscopic and supermicroscopic observations on the exciters of tuberculosis," *Klin. Wochenschrift*, Vol. 19, pp. 217–220, March 1940.
20. H. E. Morton and T. F. Anderson, "Some morphologic features of the Nichols strain of *Treponema pallidum* as revealed by the electron microscope," *Am. J. Syphilis, Gonorrhea, Venereal Diseases*, Vol. 26, pp. 565–573, September 1942.
21. H. E. Morton and T. F. Anderson, "Electron-microscopic studies of biological reactions. I. Reduction of potassium tellurite by *Corynebacterium diphtheriae*," *Proc. Soc. Exptl. Biol. Med.*, Vol. 46, pp. 272–276, 1941.
22. S. Mudd and T. F. Anderson, "Selective 'staining' for electron micrography," *J. Exptl. Med.*, Vol. 76, pp. 103–108, July 1942.
23. S. Mudd and T. F. Anderson, "Demonstration by the electron microscope of the combination of antibodies with flagellar and somatic antigens," *J. Immunol.*, Vol. 42, pp. 251–266, November 1941.
24. F. Jung, "Phenomena of degeneration in erythrocytes," *Naturwissenschaften*, Vol. 30, pp. 472–473, July 1942.
25. C. Wolpers and H. Ruska, "Structural studies of blood clotting," *Klin. Wochschrift*, Vol. 18, pp. 1077–1081; 1111–1116, August 1939.
26. A. G. Richards and T. F. Anderson, "Electron-microscope studies of insect cuticle, with a discussion of the application of electron optics to this problem," *J. Morphol.*, Vol. 71, pp. 135–183, July 1942.
27. A. G. Richards and T. F. Anderson, "Electron micrographs of insect tracheae," *J. N. Y. Entomol. Soc.*, Vol. 50, pp. 147–167, June 1942.
28. T. F. Anderson and A. G. Richards, "Electron-microscope study of some structural colors of insects," *J. Applied Phys.*, Vol. 13, pp. 748–758, December 1942.
29. A. G. Richards, T. F. Anderson, and R. T. Hance, "A microtome sectioning technique illustrated with sections of striated muscle," *Proc. Soc. Exptl. Biol. Med.*, Vol. 51, pp. 148–152, 1942.
30. A. G. Richards, H. B. Steinbach, and T. F. Anderson, "Electron-microscope studies of squid giant nerve axoplasm," *J. Cellular Comp. Physiol.*, Vol. 21, pp. 129–143, April 1943.
31a. C. E. Hall, M. A. Jakus, and F. O. Schmitt, "Electron-microscope observations of collagen," *J. Am. Chem. Soc.*, Vol. 64, p. 1234, 1942.
31b. E. F. Fullam and A. E. Gessler, "High speed microtome for the electron microscope," *J. Applied Phys.* (in press), 1945.
31c. H. C. O'Brien and G. M. McKinley, "New microtome and sectioning method for electron microscope," *Science*, Vol. 98, pp. 455–456, 1943.

32. H. O. Müller and C. W. A. Paschwaldt, "The fine structure of the test diatom *Pleurosigma angulatum* (W. Smith) from observations and stereo micrographs with the supermicroscope," *Naturwissenschaften*, Vol. 30, pp. 55–60, January *1942.*

33. *D. H. Hamly and J. H. L. Watson,* "Electron and optical microscope interpretation of the wall of *Pleurosigma angulatum*," *J. Optical Soc. Am.*, Vol. 32, pp. 433–442, August 1942.

34. M. v. Ardenne, "Electron microscopy of living substance," *Naturwissenschaften*, Vol. 29, pp. 521–523, August 1941.

35. M. v. Ardenne and H. Friedrich-Freksa, "The germination of spores of *Bacillus vulgatus* after previous imaging in the 200-kilovolt Universal Electron Microscope," *Naturwissenschaften*, Vol. 29, pp. 523–528, August 1941.

36. M. v. Ardenne, "Imaging of very fine particles, in particular molecules, with the Universal Electron Microscope," *Z. physik. Chem.*, Vol. 187, pp. 1–12, July 1940.

37. W. M. Stanley and T. F. Anderson, "Electron micrographs of protein molecules," *J. Biol. Chem.*, Vol. 146, pp. 25–30, 1942.

38. E. Husemann and H. Ruska, "Rendering visible of molecules of *p*-iodine benzoyl glycogen," *Naturwissenschaften*, Vol. 28, p. 534, August 1940.

39. B. v. Borries and G. A. Kausche, "Supermicroscopic determination of the shape and size distribution of gold colloids," *Kolloid-Z.*, Vol. 90, pp. 132–141, February 1940.

40. Columbian Carbon Company, "The particle size and shape of colloidal carbon as revealed by the electron microscope," Binney and Smith Co., distributors, New York, 1940.

41. W. R. Wiegand and W. A. Ladd, "Colloidal carbon as revealed by the electron microscope," *Rubber Age (N.Y.)*, Vol. 50, pp. 431–436, March 1942.

42. Th. Schoon and H. W. Koch, "Investigations of rubber fillers: I. Particle size and characteristic form of carbon blacks and their influence on the properties of rubber mixtures, as judged by supermicrographs," *Kautschuk*, Vol. 17, pp. 1–7, January 1941.

43. H. Mahl, "Application of the supermicroscope in colloid chemistry and metallurgy," *Jahrb. der AEG-Forschung*, Vol. 7, pp. 67–76, March 1940.

44. R. B. Barnes and C. J. Burton, "Metallic smokes as test objects in electron microscopy," *Ind. Eng. Chem.*, News Ed., Vol. 19, pp. 965–967, September 1941.

45. F. Schmieder, "Supermicroscopic study of the relation between covering power and particle size in pigments," *Kolloid-Z.*, Vol. 95, pp. 29–33, 1941.

46. B. T. Shaw and R. P. Humbert, "Electron micrographs of clay minerals," *Soil Sci. Soc. Am., Proc.*, Vol. 6, pp. 146–149, 1941.

47. C. E. Marshall, R. P. Humbert, B. T. Shaw, and O. G. Caldwell, "Studies of clay particles with the electron microscope," *Soil Sci.*, Vol. 54, pp. 149–158, 1942.

48. W. Eitel and O. E. Radczewski, "Characterization of the clay mineral montmorillonite on the supermicroscopic image," *Naturwissenschaften*, Vol. 28, pp. 628–630, September 1940.

49. R. P. Humbert and B. T. Shaw, "Studies of clay particles with the electron microscope," *Soil Sci.*, Vol. 52, pp. 481–487, 1941.

50. H. O'Daniel and O. E. Radczewski, "Electron microscopy and diffraction of highly disperse minerals on the same specimen," *Naturwissenschaften*, Vol. 28, pp. 628–630, September 1940.

51. W. Eitel and C. Schusterius, "Evaluation of supermicroscopic pictures for

the determination of the grain distribution of clays," *Naturwissenschaften*, Vol. 28, pp. 300–303, 1940.

52. W. EITEL and E. GOTTHARDT, "The stereophotogrammetric measurement of very small crystals from supermicroscopic pictures," *Naturwissenschaften*, Vol. 28, p. 367, June 1940.
53. W. EITEL, H. O. MÜLLER, and O. E. RADCZEWSKI, "Supermicroscopic study of clay minerals," *Ber. deut. keram. Gesell.*, Vol. 20, pp. 165–180, April 1939.
54. O. E. RADCZEWSKI, H. O. MÜLLER, and W. EITEL, "Supermicroscopic study of the hydration of lime," *Zement*, No. 49, pp. 1–4, 1939.
55. O. E. RADCZEWSKI, H. O. MÜLLER, and W. EITEL, "The hydration of tricalcium silicate," *Naturwissenschaften*, Vol. 27, p. 807, 1939.
56. O. E. RADCZEWSKI, H. O. MÜLLER, and W. EITEL, "The hydration of tricalcium aluminate," *Naturwissenschaften*, Vol. 27, pp. 837–838, 1939.
57. C. E. HALL and A. L. SCHOEN, "Application of the electron microscope to the study of photographic phenomena," *J. Optical Soc. Am.*, Vol. 31, pp. 281–285, April 1941.
58. L. MARTON, J. W. MCBAIN, and R. D. VOLD, "An electron-microscopy study of curd fibres of sodium laureate," *J. Am. Chem. Soc.*, Vol. 63, pp. 1990–1993, 1941.
59. M. v. ARDENNE and D. BEISCHER, "Investigation of the fine structure of high molecular substances with the Universal Electron Microscope," *Kautschuk*, Vol. 16, pp. 55–60, May 1940.
60. H. RUSKA, "On structures of cellulose fibres," *Kolloid-Z.*, Vol. 92, pp. 276–285, September 1940.
61. H. RUSKA and M. KRETSCHMER, "Supermicroscopic studies of the disintegration of cellulose fibres," *Kolloid-Z.*, Vol. 93, pp. 163–166, November 1940.
62. G. R. SEARS and E. A. KREGEL, "Application of the electron microscope to problems of the pulp, paper, and paperboard industry," *Paper Trade J.*, Vol. 114, pp. 43–49, 1942.
63. H. MAHL, "The damaging of cellulose fibres by radiation in the electron microscope," *Kolloid-Z.*, Vol. 96, pp. 7–10, 1941.
64. R. B. BARNES and C. J. BURTON, "The electron microscope," *Am. Dyestuffs Reptr.*, Vol. 31, pp. 254, 313, 1942.
65. R. B. BARNES and C. J. BURTON, "Electron microscope and cellulose," *Ind. Eng. Chem., Ind. Ed.*, Vol. 35, pp. 120–125, January 1943.
66. J. E. RUEDY, "Crystal structure and surface flow of thin evaporated antimony films," *Phys. Rev.*, Vol. 59, p. 926, 1941.
67. J. HILLIER, R. F. BAKER, and V. K. ZWORYKIN, "A diffraction adapter for the electron microscope," *J. Applied Phys.*, Vol. 13, pp. 571–577, September 1942.
68. R. G. PICARD and O. S. DUFFENDACK, "Studies on the structure of thin metallic films by means of the electron microscope," *J. Applied Phys.*, Vol. 14, pp. 291–305, June 1943.
69. R. HAEFER, "Experimental studies for testing the wave-mechanical theory of field electron emission," *Z. Physik*, Vol. 116, pp. 604–623, 1940.
70. R. D. HEIDENREICH and V. G. PECK, "Fine structure of metallic surfaces with the electron microscope," *J. Applied Phys.*, Vol. 14, pp. 23–29, January 1943.
71. V. K. ZWORYKIN and E. G. RAMBERG, "Surface studies with the electron microscope," *J. Applied Phys.*, Vol. 12, pp. 692–695, September 1941.
72. E. GÖLZ, "Supermicroscopic fine structure of glass fracture surfaces," *Z. Physik*, Vol. 120, pp. 773–777, April 1943.
73. R. F. MEHL, "The structure and rate of formation of pearlite," *Trans. Am. Soc. Metals*, pp. 813–862, December 1941.

74. R. D. HEIDENREICH, "Interpretation of electron micrographs of silica surface replicas," *J. Applied Phys.*, Vol. 14, pp. 312–320, July 1943.
75. H. MAHL, "A plastic printing process of the supermicroscopic study of metal surfaces," *Metallwirtschaft*, Vol. 19, pp. 488–491, 1940.
76. H. MAHL, "Supermicroscopic determination of the orientation of single aluminum crystals," *Metallwirtschaft*, Vol. 19, pp. 1082–1085, 1940.
77. H. MAHL, "Supermicroscopic surface representation by the replica process," *Naturwissenschaften*, Vol. 30, pp. 207–217, April 1942.
78. H. MAHL and I. N. STRANSKI, "The behavior of aluminum in crystal pattern etching," *Naturwissenschaften*, Vol. 31, pp. 12–17, January 1943.
79. B. v. BORRIES and S. JANZEN, "Imaging of finely worked technical surfaces in the supermicroscope," *Z. Ver. deut. Ing.*, Vol. 85, pp. 207–211, March 1941.

PART II. THEORETICAL BASIS OF ELECTRON OPTICS AND THE ELECTRON MICROSCOPE

CHAPTER 10

THEORETICAL BASIS OF ELECTRON OPTICS

10·1. Basic Problems of Electron-Optical Design. The primary function of electron optics is to obtain an answer to the following question: What arrangement of electrodes and field coils or magnets will cause electrons leaving one set of points to arrive at another set of points which correspond to the first set in prescribed fashion? Frequently the question is somewhat less specific. Thus, for electron guns it becomes: What electron-optical system will concentrate the electrons from a certain area of the cathode on a spot of prescribed diameter? For an electrostatic multiplier: What electrode configuration will guide electrons leaving a certain area of one electrode to a smaller area (with suitable surface-field conditions) on the next electrode? For an imaging device: What system will form a sharp and faithful image of given magnification in a position which may or may not be prescribed exactly?

The last formulation of the problem, in particular, can readily be divided into two parts: (1) the determination of a system which will form an image of given magnification in a given position and (2) the determination of a system which will give a perfect, that is, sharp and faithful, image. The first problem has an infinite number of solutions, the second, none. It may at best be possible to discover that system which, under a number of restrictive conditions, will yield the most nearly perfect image, balancing defects against each other. Even the last can be accomplished only in rare, and relatively unimportant, cases.

It is thus not surprising that in electron optics, just as in the design of light-optical instruments, methods of trial and error and interpolation play a major role. These may be purely experimental, purely mathematical, or they may combine both techniques. The purely experimental approach consists in the construction of different electrode systems or systems of field coils and pole pieces and in the observation of the results obtained as electrode potentials and field currents are varied. The systems are successively modified in directions which appear to lead to a closer approximation to the desired results. The purely mathematical approach consists in first calculating the field

distribution in an assumed electrode or coil-and-pole-piece configuration and, as a second step, in calculating the electron paths within these fields. From these paths the location, magnification, and perfection of the image, or the size and intensity distribution of the electron spot, can be determined.

Frequently, the purely experimental method proves costly and, because of the difficulty of completely eliminating extraneous influences, relatively uncertain in its conclusions. This is particularly true in the determination of lens aberrations. On the other hand, the purely mathematical method can deal effectively only with very simple configurations. A combination of the two techniques which proves valuable in many cases consists in the measurement of the field distribution, either on the original configuration or on a suitably constructed model,[1] and the subsequent calculation of the electron paths[2] and, eventually, of the aberration constants[3] of the system with the aid of the measured field. The two principal steps in the study of electron-optical systems are nearly always (1) the determination of the electric or magnetic field distribution within a given arrangement of electrodes or coils and pole pieces and (2) the determination of the electron paths within these fields.

Beyond the consideration of specific problems, it is the function of theory to explore the fundamental limitations of electron-optical systems and to establish general rules which increase the range of the conclusions that may be drawn from experimental measurements and calculations. The derivation of such rules is taken up in the succeeding chapters, which deal with suitable techniques for determining the properties of electron-optical systems and summarize the results obtained by such methods.

10·2. The Electron — Particle and Wave Aspects. Since its isolation by J. J. Thomson[4] in 1897, the electron has generally been regarded as a minute, negatively charged particle.[5] Although no exact figures can be given for its dimensions, measurements on the scattering of electrons by atomic nuclei indicate that the diameter of the electron is less than 10^{-12} centimeter. The treatment of electrons as particles subject to the laws of Newtonian mechanics has proved exceedingly fruitful. Even today it forms the basis of most of the theory of electron optics which is treated in the subsequent chapters of this book.

This relatively simple concept proved inadequate, however, to explain the behavior of electrons in atoms which is manifested in atomic line

[1] See sections 11·10 and 14·5.
[2] See Chapter 12.
[3] See Chapters 16 and 17.
[4] See reference 1.
[5] For the values of the mass and charge of the electron, as well as other physical constants, see Appendix II.

spectra. Even the introduction of restraints foreign to classical electrodynamics by the early quantum theory did not lead to a generally valid theory of atomic dynamics. The dilemma was resolved, finally, by the development of a quantum or wave mechanics which deviates in its predictions from classical mechanics only in the case of phenomena occurring on a very small (for example, atomic) scale. For large-scale phenomena, wave mechanics passes over, asymptotically, into classical mechanics.

The situation is closely analogous to that existing in light optics. Here phenomena due to light beams encountering no obstacles comparable in size to the wave length of the light can be represented adequately by ray or geometrical optics. The laws of geometrical optics are identical with those of Newton's corpuscular theory of light. Geometric optics forms the basis of optical-instrument design. However, if the light encounters finely divided gratings, minute obstacles, or very small apertures, the *diffraction patterns* observed can be explained only by reference to the wave theory of light. The application of Huygens' principle[6] shows that, for large-scale phenomena, the predictions of the wave theory of light coincide with those of geometric optics.

The characteristics of a wave motion associated with moving electrons can be derived most readily with the aid of the special theory of relativity. This was demonstrated by Louis de Broglie,[7] who first introduced the wave concept to explain the existence of stationary energy states in the hydrogen atom.

The wave properties of electrons in motion may be derived from Einstein's special theory of relativity and Planck's quantum relation. Einstein's special theory of relativity proceeds from two assumptions: (1) As in Newtonian mechanics, the laws of physics are the same for any inertial frame of reference.[8] (2) As suggested by the Michelson-Morley experiment,[9] light is propagated in vacuum with the same velocity in all such systems. To reconcile these two propositions it is necessary to drop the assumption of the existence of an absolute time applicable to all systems. Times and distances between two events measured, in a particular system, with the aid of clocks and yardsticks at rest in that system will not be identical with the times and distances between the same two events measured by clocks and yardsticks at rest in some other system moving with uniform velocity relative to the first.

[6] See, e.g., Houstoun, reference 2, p. 127.

[7] See de Broglie, reference 3.

[8] An inertial frame of reference is one in which a body, initially at rest, remains at rest unless acted upon by forces. A frame of reference which is displaced with uniform velocity relative to an inertial frame is also an inertial frame.

[9] See, e.g., Miller, reference 4.

Consider (Fig. 10·1) two frames of reference with the coordinates (x,y,z,t) and (x_o,y_o,z_o,t_o), with the x- and x_o-axes coincident, and the y- and z-axes parallel to the y_o- and z_o-axes, respectively. Furthermore, let the origin of the frame of reference (x_o,y_o,z_o,t_o) move with the uniform

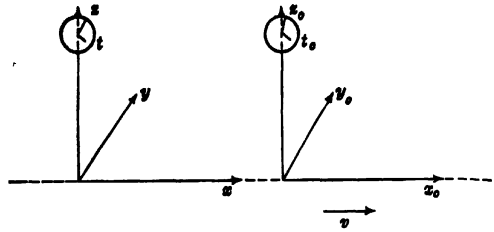


FIG. 10·1. Frames of Reference for the Derivation of the Lorentz Transformation.

velocity v in the x-direction relative to the "fixed" frame (x,y,z,t), as measured by an observer in the fixed system. Finally, imagine the several clocks to be set so that $t = t_o = 0$ at the moment the two systems coincide.

In Newtonian mechanics the relation between the coordinates of an event in the two frames would be:

$$x_o = x - vt \qquad y_o = y \qquad z_o = z \qquad t_o = t \qquad [10·1]$$

This conflicts with the second requirement, which demands, for example, that the equation for the wave front of a light wave proceeding from the origin at the instant $t = 0$ should have the same form in the two systems. Thus the equation

$$x_o^2 + y_o^2 + z_o^2 = c^2t_o^2 \qquad [10·2]$$

where c is the velocity of light, must pass over into

$$x^2 + y^2 + z^2 = c^2t^2 \qquad [10·3]$$

This requirement is satisfied with retention of the linear relation between the two sets of coordinates if the Eqs. 10·1 are replaced with the so-called Lorentz transformation equations:

$$x_o = \frac{x - vt}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \qquad y_o = y$$

$$z_o = z \qquad t_o = \frac{t - \frac{vx}{c^2}}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \qquad [10·4]$$

It follows readily from these equations that no mechanical velocity can exceed that of light, c. Further conclusions of the special theory of relativity, arising from the requirement that the law of conservation of momentum be fulfilled in both coordinate systems and from the integration of Newton's (duly interpreted) second law, are:

1. The variation of inertial mass with velocity:

$$m = \frac{m_o}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}} \qquad [10\cdot5]$$

2. The proportionality of the mass and the total energy of a material body:

$$E = mc^2 = \frac{m_o c^2}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}} = \text{kinetic energy} + m_o c^2 \qquad [10\cdot6]$$

Here m_o is the rest mass of the body and $m_o c^2$ the energy associated therewith. If the expression for the kinetic energy given by Eq. 10·6 is expanded in terms of v^2/c^2, the first term becomes identical with the classical value $m_o v^2/2$:

$$\text{Kinetic energy} = \frac{m_o v^2}{2}\left(1 + \frac{3v^2}{4c^2} + \frac{5v^4}{8c^4} + \cdots\right) \qquad [10\cdot7]$$

Returning now to the wave theory of matter, assume that an oscillation is associated with the stationary electron, represented by

$$A \sin 2\pi\nu_o t_o \qquad [10\cdot8]$$

where the frequency ν_o is fixed, as for a light quantum or photon, by the relation

$$E_o = h\nu_o \qquad \nu_o = \frac{E_o}{h} = \frac{m_o c^2}{h} \qquad [10\cdot9]$$

In a system relative to which the electron moves with a velocity v the oscillation associated with the electron will take on the form of a progressive wave, given, according to Eq. 10·4, by

$$A \sin 2\pi\nu\left(t - \frac{vx}{c^2}\right) = A \sin 2\pi\nu\left(t - \frac{x}{u}\right) = A \sin 2\pi\left(\nu t - \frac{x}{\lambda}\right) \qquad [10\cdot10]$$

Here u, the wave velocity, is defined in the usual manner as the velocity with which a point of constant phase (for example, a crest of the wave) moves forward and λ, the wave length, as the distance between two

successive crests at any instant. Also, by Eqs. 10·4 and 10·9:

$$\nu = \frac{\nu_0}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}} = \frac{m_0 c^2}{h\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}} \qquad [10\cdot11]$$

$$u = \frac{c^2}{v} \qquad [10\cdot12]$$

and

$$\lambda = \frac{u}{\nu} = \frac{h}{mv} = \frac{h\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}}{m_0 v} \qquad [10\cdot13]$$

It remains to indicate how the matter wave, propagated with a velocity in excess of that of light, is prevented from becoming detached from the material particle, progressing with a velocity v. In other words, how is the particle velocity reflected in the wave phenomenon?

The answer becomes apparent when an interpretation of the significance of the matter wave is attempted. For a light wave, the square of the amplitude of the wave indicates the amount of light energy per unit volume in the wave or the number of light quanta or photons which may be expected to be found in unit volume at the point in question. It is thus reasonable to identify the square of the amplitude of the matter wave with the concentration of energy or the probability of locating a particle in unit volume at the point considered.

With this interpretation it becomes evident that the plane electron wave in Eq. 10·10 does not represent a single localized material particle such as an electron, but a uniform stream of particles traveling in the x-direction, filling all space. To represent a localized electron it is necessary to construct a wave group, consisting of the superposition of many wave components differing slightly in direction of propagation, frequency, and, hence, wave velocity:

$$\int_{-\delta\nu}^{\delta\nu}\int_{-\delta m}^{\delta m}\int_{-\delta n}^{\delta n} A(\nu + \epsilon, m, n) \cdot \sin 2\pi(\nu + \epsilon)\cdot\left[t - \frac{x + my + nz}{u(\nu + \epsilon)}\right] d\epsilon\, dm\, dn \qquad [10\cdot14]$$

where δm and δn, the maximum values of the direction cosines m and n, are small compared to 1, and $\delta\nu$ is small compared to ν. This may be regarded as an illustration of one of the fundamental principles of the

newer quantum theory: the uncertainty relation of Heisenberg.[10] According to this, the precision with which the location of an object can be determined is inversely proportional to the accuracy with which its momentum (or velocity) may be known, the magnitude of the uncertainties in the coordinate and velocity components being related to Planck's constant h in the following manner:

$$\Delta x \cdot \Delta(mv_x) = \Delta y \cdot \Delta(mv_y) = \Delta z \cdot \Delta(mv_z) = \frac{h}{2\pi} \qquad [10\cdot15]$$

In the case of the plane wave, the momentum (and hence the frequency and wave length) of the particle is known exactly and its position is completely indefinite:

$$\Delta(mv_x) = \Delta(mv_y) = \Delta(mv_z) = 0 \qquad \Delta x = \Delta y = \Delta z = \infty$$

The manner in which a group of wave components with different wave lengths may cooperate to form a single wave group is indicated in Fig. 10·2. It is seen that the peak occurs where the phase ψ of the

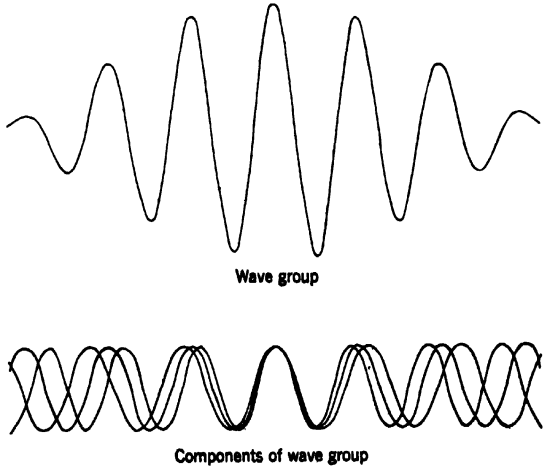


FIG. 10·2. The Formation of a Wave Group as Resultant of Wave Trains with Slightly Differing Wave Length.

several components is the same, that is, where $d\psi/dv = 0$. This condition also determines the velocity of propagation of the peak and, hence, the wave group, v_g. Thus for wave components propagated in

[10] See reference 5.

the x-direction ($m = n = 0$):

$$\frac{d\psi}{d\nu} = 2\pi\left\{t - x\frac{d}{d\nu}\left(\frac{\nu}{u}\right)\right\} = 0 \quad \text{and} \quad v_g = \frac{x}{t} = \frac{1}{\dfrac{d\left(\dfrac{\nu}{u}\right)}{d\nu}} \qquad [10\cdot16]$$

However,

$$\frac{d}{d\nu}\left(\frac{\nu}{u}\right) = \frac{d}{d\nu}\left\{\frac{\nu}{c}\left[1 - \frac{m_o^2c^4}{h^2\nu^2}\right]^{\frac{1}{2}}\right\} = \frac{1}{c\left[1 - \frac{m_o^2c^4}{h^2\nu^2}\right]^{\frac{1}{2}}} = \frac{1}{v} \qquad [10\cdot17]$$

so that

$$v_g = v$$

Thus the velocity of the wave group is identical with the particle velocity, as must be the case if the wave intensity is interpreted as a measure of probability of the presence of the particle.

The important conclusion of these considerations is that the analogy between moving particles (electrons) and light goes beyond that between particle mechanics and geometrical optics. The parallelism between electrons and photons is supplemented by that between electron waves and light waves, the relationship between particle and wave being identical in the two cases. Thus not only the knowledge gathered in the field of geometrical light optics becomes applicable to electron optics, but also that obtained in the field of wave optics, relating to diffraction and interference. According to Eq. 10·13 the wave length of an electron beam is given by $\lambda = h/(mv)$. If the electric potential ϕ of the field through which the beam passes is fixed at zero wherever the velocity of the electrons is reduced to zero, their kinetic energy becomes $e\phi$, so that

$$m_oc^2\left(\frac{1}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}} - 1\right) = e\phi$$

$$\frac{v}{c} = \frac{m_oc^2}{e\phi + m_oc^2}\left[\left(\frac{e\phi}{m_oc^2}\right)^2 + \frac{2e\phi}{m_oc^2}\right]^{\frac{1}{2}} \qquad [10\cdot18]$$

and

$$\frac{mv}{c} = m_o\left[\left(\frac{e\phi}{m_oc^2}\right)^2 + \frac{2e\phi}{m_oc^2}\right]^{\frac{1}{2}} \qquad [10\cdot19]$$

The expression for the wave length λ thus becomes

$$\lambda = \frac{h}{m_o c\left[\frac{2e\phi}{m_o c^2} + \left(\frac{e\phi}{m_o c^2}\right)^2\right]^{\frac{1}{2}}} \qquad [10\cdot20]$$

If the wave length is expressed in angstrom units and ϕ in volts, this becomes

$$\lambda = \frac{12.26}{[\phi + 0.978\cdot10^{-6}\phi^2]^{\frac{1}{2}}} \text{ A.U.} \qquad [10\cdot21]$$

For relatively low voltages the wave length is thus inversely proportional to the square root of the accelerating voltage of the electrons — $\lambda \approx 12.3/\phi^{\frac{1}{2}}$ A.U. If the accelerating voltage is smaller than 20,000 volts, this approximate formula is in error by less than 1 per cent. It is seen that the wave lengths associated with electron beams are exceedingly short. For 50,000 volts Eq. 10·21 yields $\lambda = 0.0535$ A.U. As has been brought out in section 3·1, the fact that the wave lengths of electron beams are so much shorter than those of light beams ($\sim$5000 A.U.) and that diffraction effects are correspondingly less pronounced forms the basis of the superior resolving power of electron microscopes.

10·3. Electron Optics of Electrostatic Fields. The fundamental law of geometrical light optics is Fermat's law. This states that among a number of possible adjoining paths between two points A and B, a light ray will take that path which reduces the time of transit to a minimum, or, more generally, gives it a stationary value. In other words, the rate of change of the time of transit with the magnitude of any deviation from the path considered vanishes for the actual light path. In conventional symbols this is expressed by the vanishing of the *variation* of the transit time:

$$\delta \int_A^B \frac{ds}{u} = \delta \int_A^B \frac{n}{c} ds = 0 \qquad [10\cdot22]$$

Here u is the velocity of propagation of the light wave, equal to the ratio of the velocity of light in vacuum, $c = 3\cdot10^{10}$ centimeters per second, and the index of refraction n of the medium through which the path passes; ds is an element of that path. In ordinary, isotropic media n is a function of position only, whereas in anisotropic crystalline media it depends on the direction of travel of the light wave as well. The familiar laws of reflection and refraction can readily be shown to follow from Fermat's theorem. Since the product of path length and

index of refraction is also known as optical distance, Fermat's law may be stated in the form that, for the actual ray path, the optical distance between two given terminal points assumes a stationary value.

A closely corresponding principle governs the motion of material particles in conservative force fields.[11] This is the law of least action or

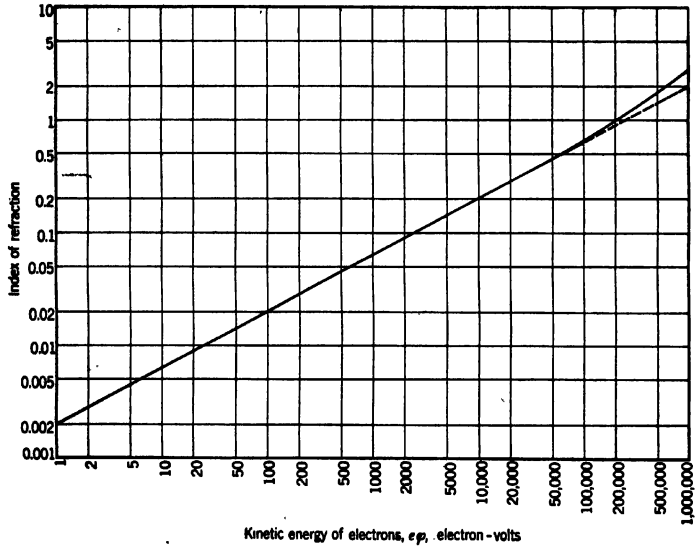


FIG. 10·3. Variation of Electron Index of Refraction with Accelerating Voltage ϕ.

principle of Maupertuis. According to it a particle of given initial kinetic energy will select that path among neighboring paths with the same terminal points A and B for which the action, or the integral of the momentum over the path, assumes a stationary value:

$$\delta \int_A^B mv\, ds = 0 \qquad [10\cdot23]$$

This goes over into a form identical with that of Fermat's law if the momentum — which for particles of the same energy in a conservative force field is a function of position only — is defined as the index of refraction of the field. Since the addition of constant factors does not

[11] A conservative force is here one which may be derived from the potential energy of the particle in the field. Both gravitational and electrostatic fields are, in this sense, conservative force fields; magnetic fields are not.

change the significance of the variation principles, a more suitable, dimensionless definition of the refractive index is

$$n = \frac{mv}{m_o c} \qquad [10\cdot24]$$

where m_o is the rest mass of the particle and c the velocity of light in vacuum. For electrons in an electrostatic field this becomes, by Eq. 10·19,

$$\begin{aligned} n &= \left(\frac{2e}{m_o c^2}\right)^{1/2} \left(\phi + \frac{e}{2m_o c^2}\phi^2\right)^{1/2} \\ &= 1.978\cdot10^{-3}(\phi + 0.978\cdot10^{-6}\phi^2)^{1/2} \end{aligned} \qquad [10\cdot25]$$

Figure 10·3 shows a plot of the index of refraction as function of the potential ϕ, whose zero level has been adjusted to correspond to the point at which the kinetic energy of the electron vanishes. It is seen that, for moderate accelerating voltages, the electron-optical index of refraction is simply proportional to the square root of the potential.

The variation principle, Eq. 10·23, necessarily leads to the same families of permitted paths as the more familiar equations of motion. This is demonstrated in detail in section 12·1.

10·4. Electron Optics of Magnetic Fields. In the presence of a magnetic field Maupertuis' principle ceases to be applicable, since the force acting on a charged particle cannot be expressed as a derivative of the potential energy of the particle in the field. In fact, the magnetic field exerts a force at right angles to the direction of motion of the particle, and hence has no effect on the kinetic energy of the particle. Even so, as was first demonstrated by Schwarzschild,[12] a variation principle may be set up which governs the motion of charged particles in magnetic fields and is analogous to Eq. 10·23. This may be written

$$\delta \int_A^B \left(mv - \frac{eA}{c}\cos\chi\right) ds = 0 \qquad [10\cdot26]$$

Here A is the magnitude of the vector potential of the magnetic field and χ the angle between the direction of the vector potential and the element of path ds at the point in question. The vector potential is related to the field $\mathbf{h}$ by the vector equation

$$\mathbf{h} = \text{curl } \mathbf{A} \qquad [10\cdot27]$$

its variation in space being fixed, except for an immaterial constant additive vector, by the simultaneous requirement div $\mathbf{A} = 0$.

[12] See reference 6.

For the special case of an axially symmetric magnetic field, **A** is normal both to the axis and to the radius vector from the axis to the point of reference. If h_z is the axial component of the field and r_o the radius vector to the point of reference,

$$A = \frac{\int_o^{r_o} h_z \cdot 2\pi r \, dr}{2\pi r_o} = \frac{N}{2\pi r_o} \qquad [10\cdot28]$$

Here the integral is to be carried out over the area normal to the axis bounded by a circle of radius r_o through the point of reference. N denotes the total magnetic flux through this area.

The index of refraction for an electron in the presence of a magnetic field is, by Eq. 10·26,

$$\begin{aligned} n &= \frac{mv}{m_o c} - \frac{e}{m_o c^2} A \cos \chi \\ &= 1.978 \cdot 10^{-3} (\phi + 0.978 \cdot 10^{-6} \phi^2)^{\frac{1}{2}} - 0.587 \cdot 10^{-3} A \cos \chi \end{aligned} \qquad [10\cdot29]$$

Here ϕ is measured in volts and A in gauss-centimeters. The peculiar property of this refractive index is that it depends not only on the position of the electron, but also on its direction of motion. This property is shared by light propagated in anisotropic — uniaxial or biaxial — crystals. Apart from this, the behavior of electrons in magnetic fields and of light in crystals differs so profoundly that crystal optics does not contribute materially to an understanding of magnetic electron optics.

10·5. Intensity Relations in Electron Beams and Electron Images. In many applications of electron optics it is desired to obtain an electron image or electron spot of very high intensity. In such cases it is of some importance to recognize the limitations which, regardless of the optical system employed, are fixed by the specific emission of the source, the convergence of the electron paths at the image or spot, and the initial and final kinetic energies of the electrons. This fundamental limitation may be derived with the aid of the second law of thermodynamics. This may be stated in this form: If two bodies are left to themselves, the colder of the two cannot transfer more heat to the warmer than the warmer transfers to the colder; if the two bodies are at the same temperature, the transfer of heat in the two directions must, consequently, be just equal.

Consider (Fig. 10·4) two black bodies 1 and 2 at the same temperature and immersed in media of refractive index n_1 and n_2. These media may be imagined to be separated from the emitting surfaces by narrow gaps, in which the index of refraction is assumed to be unity,

as shown in the figure. Let the two bodies communicate by means of an optical system which restricts the aperture of the transmitted pencils to α_1 and α_2, respectively. Let the emission per unit area at the given common temperature be j. Then the total energy transmitted by each

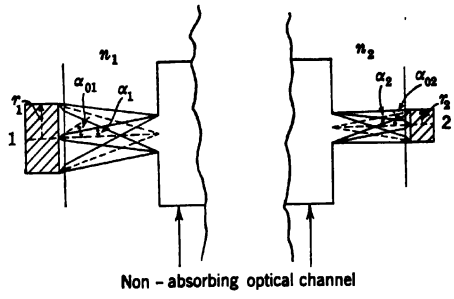


FIG. 10·4. The Derivation of the Theorem of Clausius or Helmholtz-Lagrange.

side through the optical system (which is assumed to absorb no energy) to the other must be the same. If the radii of the two (circular) emitting areas are r_1 and r_2, respectively, this leads to

$$\pi r_1^2 \cdot 2\pi \int_0^{\alpha_{01}} \frac{j}{\pi} \cos\theta \sin\theta \, d\theta = \pi r_2^2 \cdot 2\pi \int_0^{\alpha_{02}} \frac{j}{\pi} \cos\theta \sin\theta \, d\theta$$

since the surfaces obey Lambert's law, so that the emission per unit area per unit solid angle in the θ-direction is given by $(j/\pi) \cos\theta$. If the integration is carried out,

$$r_1^2 \sin^2 \alpha_{01} = r_2^2 \sin^2 \alpha_{02}$$

Here α_{01} and α_{02} are the angular apertures of the transmitted pencils in the gaps just in front of the emitting surfaces. Since, by Snell's law,

$$\sin \alpha_{01} = n_1 \sin \alpha_1 \qquad \sin \alpha_{02} = n_2 \sin \alpha_2$$

$$n_1 r_1 \sin \alpha_1 = n_2 r_2 \sin \alpha_2 \qquad [10\cdot30]$$

In the special case that an object at 1 is imaged by an axially symmetric optical system at 2, Eq. 10·30 establishes a general relation between the magnification (r_2/r_1) and the ratio of the sines of the convergence angles α_1 and α_2. This relation is commonly referred to as the theorem of Helmholtz-Lagrange, although it was first derived by Clausius.

More generally, Eq. 10·30 establishes an upper limit for the light (or electron) concentration obtainable in a beam derived from a given

source. If there is no loss of light energy or electron current in the optical system, and i represents the total emission of the source and j_1 its specific emission, the light flux or current per unit area j_2 in the spot or image is given by the relationship

$$\frac{j_2}{j_1} = \frac{i}{\pi r_2^2} \frac{\pi r_1^2}{i} = \frac{n_2^2}{n_1^2} \frac{\sin^2 \alpha_2}{\sin^2 \alpha_1} \qquad [10\cdot31]$$

Assume, now, that the source is a thermionic cathode at a temperature T. The peak of the energy distribution of the thermionic electrons then lies at the energy of emission kT, where k is Boltzmann's constant. Since the energy of arrival at the spot or image is eV_2, V_2 being the anode potential, and the aperture angle of the pencils leaving the cathode is $\pi/2$,

$$\frac{j_2}{j_1} = \frac{eV_2}{kT} \sin^2 \alpha_2 \qquad [10\cdot32]$$

A more exact derivation of this ratio, taking account of the Maxwellian distribution of the electrons leaving the cathode, has been given by D. B. Langmuir.[13] Langmuir finds

$$\frac{j_2}{j_1} = \left[\frac{eV_2}{kT} + 1\right] \sin^2 \alpha_2 = \left[\frac{11{,}600V_2}{T} + 1\right] \sin^2 \alpha_2 \qquad [10\cdot33]$$

which is identical with Eq. 10·32 provided that $eV_2 \gg kT$, which is nearly always fulfilled.

It will be noted that the basic equation (10·30) was derived making use of the properties of thermal or light radiations and is subsequently applied to electron-optical systems. This is justified by the formal identity of the laws governing light rays and electron rays (Eqs. 10·22, 10·25, and 10·29). It is necessary, however, that the electron source should obey Lambert's law if the upper limits for the current density here derived are to be valid. This is fulfilled for a thermionic cathode. Normally deviations from Lambert's law will result in but minor changes in the maximum current density obtainable.

REFERENCES

1. J. J. Thomson, "Cathode rays," *Phil. Mag.*, Vol. 4, pp. 293–316, October 1897.
2. R. A. Houstoun, *A Treatise on Light*, Longmans, Green and Co., London, 1924.
3. L. de Broglie, "Thèse de doctorat," Masson, Paris, 1924.
4. D. C. Miller, "Absolute motion of Solar system and orbital motion of the Earth by the ether-drift experiment," *Rev. Modern Phys.*, Vol. 5, pp. 203–242, July 1933.

[13] See reference 7.

5. W. HEISENBERG, "The actual content of quantum-theoretical kinematics and mechanics," *Z. Physik*, Vol. 43, pp. 172–198, 1927.
6. K. SCHWARZSCHILD, "Two forms of the principle of least action in the theory of electrons," *Gesell. Wiss. Goettingen, Nachr. Math.-Phys. Klasse*, Vol. 3, pp. 126–131, 1903.
7. D. B. LANGMUIR, "Limitations of cathode-ray tubes," *Proc. Inst. Radio Engrs.*, Vol. 25, pp. 977–991, August 1937.

CHAPTER 11

DETERMINATION OF POTENTIAL DISTRIBUTION

11·1. The Laplace Equation. Within any region of space bounded by electrodes at different voltages, an electric field will exist. In general this field will not be uniform but will vary from point to point. As has already been pointed out, the first step in the solution of most electron-optical problems is the determination of the distribution of such a field or of the potential from which it is derived.

Although the field within a given region may vary from point to point, its value at one point is not unrelated to that at any other, because the field as a whole must satisfy the condition of zero divergence, namely:

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0 \qquad [11\cdot1]$$

This relation is merely a statement of the continuity of the electrostatic field lines. Furthermore, such a field must be unrotational, that is, it must satisfy the differential equations

$$\left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}\right) = \left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y}\right) = \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z}\right) = 0 \qquad [11\cdot2]$$

Finally the field must have zero tangential components along all the conducting surfaces bounding the region.

Any field meeting these conditions can be derived from a potential ϕ, a potential which itself satisfies the differential equation

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0 \qquad [11\cdot3]$$

This relationship is known as the Laplace equation. Of course, the potential must also satisfy the boundary conditions, that is, the potential distribution must be such that the electrodes themselves lie on equipotential surfaces which have the corresponding potential. It can be shown that a potential distribution which satisfies Laplace's equation and the boundary conditions is unique, so that any method of solution leading to a distribution fulfilling these conditions must necessarily give the desired potential distribution.

Where the field is required, it may be found from the potential by

differentiating with respect to the various coordinates. Thus

$$E_x = -\frac{\partial \phi}{\partial x} \qquad E_y = -\frac{\partial \phi}{\partial y} \qquad E_z = -\frac{\partial \phi}{\partial z} \qquad [11\cdot 3a]$$

The determination of the potential distribution for a general three-dimensional set of boundary conditions is usually extremely difficult. Fortunately most of the configurations encountered in practical electron-optical problems have symmetry properties which to some extent facilitate the solution. The two most commonly encountered types of symmetry are plane (*two-dimensional*) and axial symmetry. These two classes of electrode configurations warrant examination in some detail.

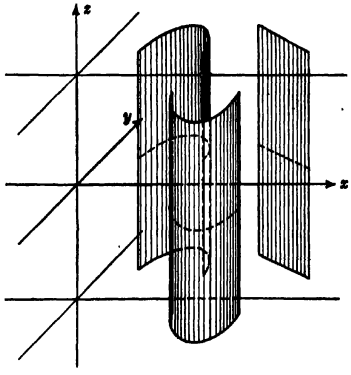


FIG. 11·1. A "Two-Dimensional" System of Electrodes.

11·2. Two-Dimensional Systems. The two-dimensional electron-optical system finds frequent practical employment in such devices as secondary-emission multipliers, deflecting systems, and collimating slit systems. In this system the electrodes may be thought of as being generated by the motion of straight lines parallel to an axis, for example, the z-axis, along paths in the xy-plane (Fig. 11·1). For the purpose of analysis, the generatrix can be assumed to be infinite, an approximation which is justifiable for most practical calculations. It is evident that, in such a system, any plane parallel to the xy-plane will be a plane of mirror symmetry; hence the potential does not depend on the z-coordinate. The Laplace equation thus becomes

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \qquad [11\cdot 4]$$

since $\dfrac{\partial^2 \phi}{\partial z^2}$ is zero everywhere. In polar coordinates, the equation may be written

$$\frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial \phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \phi}{\partial \theta^2} = 0 \qquad [11\cdot 5]$$

Here $r = (x^2 + y^2)^{1/2}$ is the radius vector and $\theta = \arctan(y/x)$ the azimuth of the point of reference.

One approach to the solution of these equations is through the fact that they are the real parts of analytic functions of the complex variable

$$w = x + iy = r \cdot e^{i\theta} \qquad [11\cdot6]$$

where i is the imaginary quantity $(-1)^{1/2}$ and e the Napierian base $(2.718\cdots)$. This can be shown readily in the following way. If

$$F(w) = \phi(x,y) + i\cdot\psi(x,y) \qquad [11\cdot7]$$

then

$$\frac{\partial F}{\partial x} = \frac{dF}{dw}\frac{\partial w}{\partial x} = \frac{dF}{dw} = \frac{\partial \phi}{\partial x} + i\frac{\partial \psi}{\partial x}$$

$$\frac{\partial F}{\partial y} = \frac{dF}{dw}\frac{\partial w}{\partial y} = i\frac{dF}{dw} = \frac{\partial \phi}{\partial y} + i\frac{\partial \psi}{\partial y}$$

so that

$$\frac{\partial \phi}{\partial x} + i\frac{\partial \psi}{\partial x} = -i\frac{\partial \phi}{\partial y} + \frac{\partial \psi}{\partial y}$$

or

$$\frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y} \qquad \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x} \qquad [11\cdot8]$$

Differentiating these final two equations by x and y, respectively, leads to

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial^2 \psi}{\partial x \partial y} \qquad \frac{\partial^2 \phi}{\partial y^2} = -\frac{\partial^2 \psi}{\partial y \partial x}$$

Since the order of differentiation makes no difference, $\phi(x,y)$ is seen to fulfil, quite generally, the Laplace equation (Eq. 11·4):

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0$$

It is interesting to note (Fig. 11·2) that the two families of curves $\phi(x,y) = \text{const}$ and $\psi(x,y) = \text{const}$ are orthogonal, as can be seen from the fact that when the first of Eqs. 11·8 is divided by the second, yielding

$$\frac{\dfrac{\partial \phi}{\partial x}}{\dfrac{\partial \phi}{\partial y}} = -\frac{\dfrac{\partial \psi}{\partial y}}{\dfrac{\partial \psi}{\partial x}}$$

$$\left(\frac{dy}{dx}\right)_{\phi=\text{const}} = -\frac{1}{\left(\dfrac{dy}{dx}\right)_{\psi=\text{const}}} \qquad [11\cdot9]$$

the condition for the orthogonality of two intersecting curves ϕ = const and ψ = const is fulfilled. This means that if ϕ = const represents a family of equipotential surfaces, ψ = const represents the electrostatic

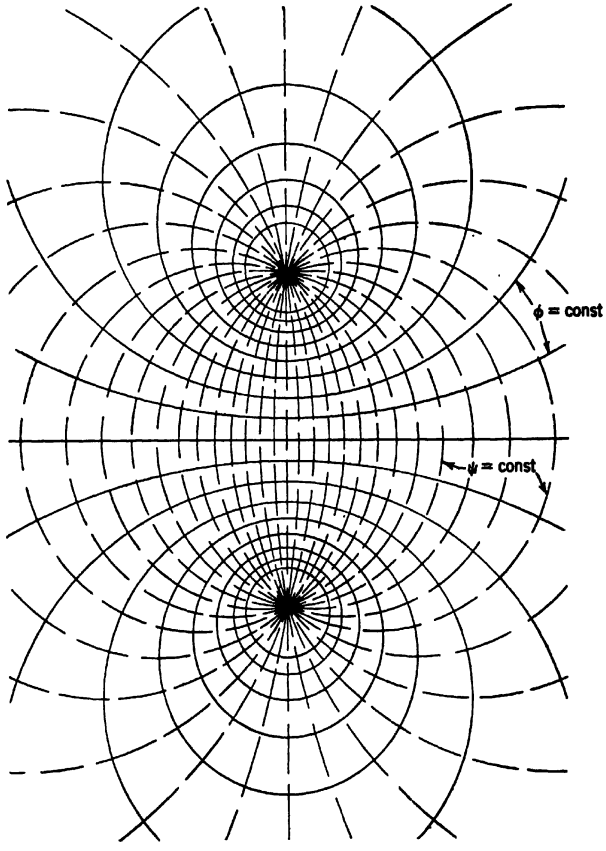


FIG. 11·2. Equipotentials (ϕ = const) and Field Lines (ψ = const) as Orthogonal Systems (Two Line Sources of Opposite Sign).

lines of force. Similarly, if ψ = const corresponds to the equipotentials, implying that Eq. 11·7 has been multiplied by the factor $-i$, ϕ = const will describe the lines of force.

A second powerful approach to the solution of the two-dimensional Laplace equation is found in the method of the *separation of variables*.

For this, the potential function is set equal to the product of two functions, each dependent on a single variable, expressed symbolically as follows:

$$\phi(x,y) = F(x)\cdot G(y) \qquad [11\cdot10]$$

When this expression is substituted in Eq. 11·4, the equation becomes

$$G(y)\frac{d^2F(x)}{dx^2} + F(x)\frac{d^2G(y)}{dy^2} = 0$$

and, upon dividing by $F(x)\cdot G(y)$, yields the following relation:

$$\frac{1}{F(x)}\frac{d^2F(x)}{dx^2} + \frac{1}{G(y)}\frac{d^2G(y)}{dy^2} = 0 \qquad [11\cdot11]$$

Since each of the terms in Eq. 11·11 depends upon one coordinate alone, and is entirely independent of the other coordinate, the equation can be satisfied for all coordinate values only if each term individually is equal to a constant. This constant is usually known as the *separation parameter* and will be denoted by k^2. Therefore Eq. 11·11 may be divided into the following two equations:

$$\frac{1}{F(x)}\frac{d^2F(x)}{dx^2} = k^2 \qquad \frac{1}{G(y)}\frac{d^2G(y)}{dy^2} = -k^2 \qquad [11\cdot12]$$

These equations are solved by the following two functions:

$$F(x) = A\sinh kx + B\cosh kx$$
$$G(y) = C\sin ky + D\cos ky \qquad [11\cdot13]$$

where k may be real or imaginary.

A general solution of the Laplace equation is found by forming the sum of all products $F(x)\cdot G(y)$ with different parameters k and arbitrary coefficients:

$$\phi(x,y) = \sum_k (A_k \sinh kx + B_k \cosh kx)(\sin ky + D_k \cos ky) \qquad [11\cdot14]$$

It will be noted that C_k has been set equal to unity. This can be done without loss of generality, since A_k and B_k are arbitrary constants. The three sets of coefficients, A_k, B_k, and D_k, appearing in this summation are evaluated from the boundary conditions of the specific problem.

Separation of the variables can, of course, also be used in the solution of the two-dimensional Laplace equation in polar coordinates (Eq. 11·5), the choice of coordinates being dictated by the relative simplicity of the boundary conditions. For this it is assumed that the potential function can be separated as follows:

$$\phi(r,\theta) = f(r)\cdot g(\theta) \qquad [11\cdot15]$$

This leads to two differential equations related through the separation parameter k. The general solution can be expressed as the summation

$$\phi(r,\theta) = \sum_k \left(A_k r^k + \frac{B_k}{r^k}\right)(\sin k\theta + D_k \cos k\theta) \qquad [11\cdot16]$$

Here, again, one coefficient (that of $\sin k\theta$) has been set equal to unity, since this involves no loss of generality.

Two other general methods which are of great value when the electrode configurations are too complex to handle analytically are the Liebmann net procedure and electrolytic potential plotting. The discussion of these very general methods of determining potential distributions will be deferred until sections 11·9 and 11·10.

11·3. Potential Distribution by Means of Complex Functions. Certain potential distributions, which can be found readily by direct application of complex functions as outlined in the preceding section, are of such a form as to have considerable practical importance. Since, in addition to their intrinsic value, they will serve to illustrate the application of this method, a few will be derived below.

a. Parallel Planes. If the complex function

$$F(w) = Aw + B \qquad [11\cdot17]$$

is chosen, it gives rise to the orthogonal family

$$\phi(x,y) = \text{const } x + k \qquad [11\cdot18]$$

$$\psi(x,y) = \text{const } y \qquad [11\cdot19]$$

It is evident that the equipotentials consist of a family of parallel planes. Any pair of these may be chosen as electrodes of the system. Therefore, if the electrodes consist of parallel planes at $x = x_1$ and $x = x_2$, having potentials $\phi = V_1$ and $\phi = V_2$, respectively, as illustrated in Fig. 11·3*a*, the potential distribution and field function are given by

$$\phi(x,y) = \frac{V_2 - V_1}{x_2 - x_1}x - \frac{V_2x_1 - V_1x_2}{x_2 - x_1} \qquad [11\cdot20]$$

$$\psi(x,y) = \frac{V_2 - V_1}{x_2 - x_1}y \qquad [11\cdot21]$$

b. Coaxial Cylinders. If the function assumed is $F(w) = A \log w + B$, then, since in polar coordinates w may be written

$$w = r\cdot e^{i\theta} \qquad [11\cdot22]$$

the complex function becomes

$$F = A\ (\log r + i\theta) + B \qquad [11\cdot23]$$

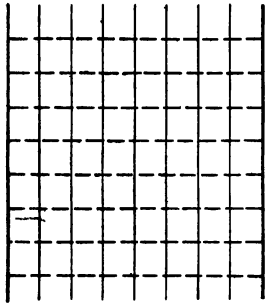


FIG. 11·3(*a*). Parallel Planes.

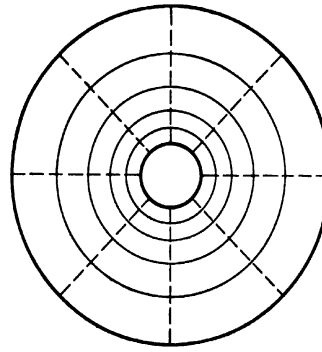


FIG. 11·3(*b*). Coaxial Cylinders.

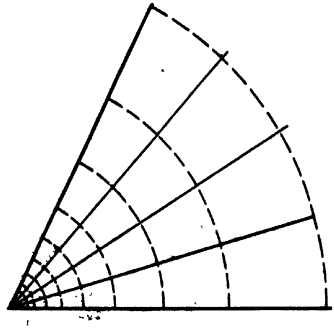


FIG. 11·3(*c*). Intersecting Planes.

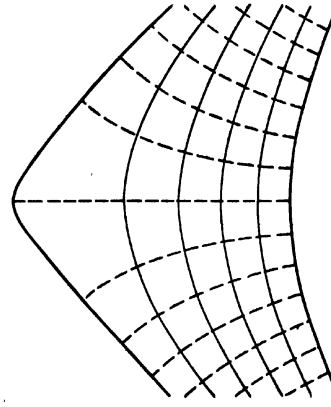


FIG. 11·3(*d*). Rectangular Hyperbolic Cylinders.

FIG. 11·3. Equipotentials and Field Lines for Specific Two-Dimensional Electrode Systems (——— Equipotentials, – – – – Field Lines).

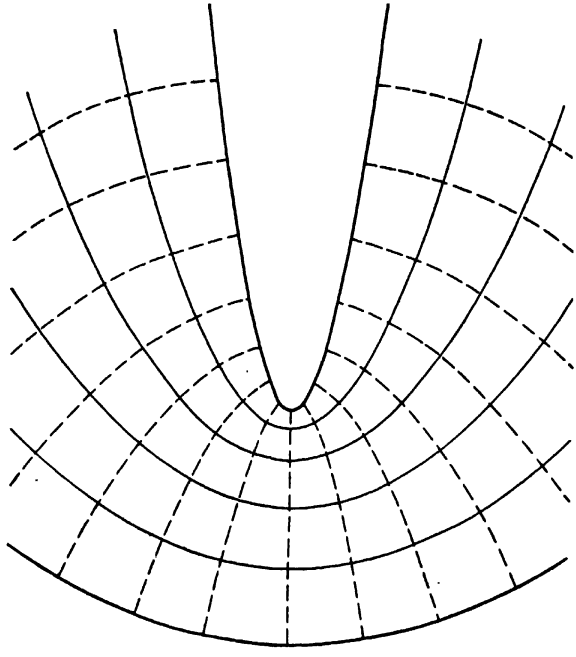


FIG. 11·3(e). Confocal Parabolic Cylinders.

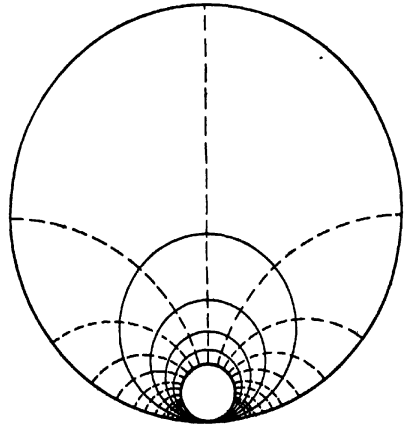


FIG. 11·3(f). Tangent Circular Cylinders.

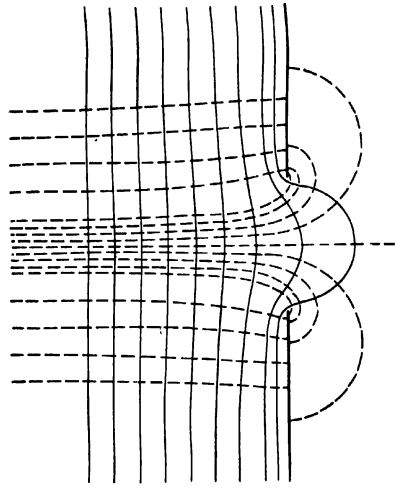


FIG. 11·3(g). Slit ($E_2 = 0$).

Hence the equipotential surfaces are a family of coaxial cylinders whereas the field lines are straight radial lines. When, as illustrated in Fig. 11·3*b*, the electrodes are cylinders of radius r_2 and r_1, at potentials V_2 and V_1, respectively, the potential function becomes

$$\phi(r) = \frac{V_2 - V_1}{\log \frac{r_2}{r_1}} \log \frac{r}{r_1} + V_1 \qquad [11\cdot24]$$

Similarly, the electrostatic lines of force are given by

$$\psi(\theta) = \frac{V_2 - V_1}{\log \frac{r_2}{r_1}} \theta \qquad [11\cdot25]$$

c. Intersecting Planes. The potential distribution of a system made up of two intersecting planes, as illustrated in Fig. 11·3*c*, can be obtained from case *b* by the transformation mentioned in the preceding section, whereby the roles of the real and imaginary parts of the function are reversed. By multiplying the preceding function by i it becomes

$$F(w) = Ai \log w + iB \qquad [11\cdot26]$$

or

$$F(w) = A(i \log r - \theta) + iB \qquad [11\cdot27]$$

The field lines are now concentric circles (or parts thereof) whereas the equipotential surfaces are radial planes. Where the two electrodes make angles θ_2 and θ_1 with an arbitrary reference plane, and are at potentials V_2 and V_1, respectively, the potential distribution is

$$\phi(\theta) = \frac{V_2 - V_1}{\theta_2 - \theta_1}\theta + \frac{V_1\theta_2 - V_2\theta_1}{\theta_2 - \theta_1} \qquad [11\cdot28]$$

and the field lines are represented by

$$\psi(r) = -\frac{V_2 - V_1}{\theta_2 - \theta_1}\log r \qquad [11\cdot29]$$

A number of other useful distributions are given below. In the configurations listed, with the exception of case *g*, V_2 and V_1 are the potentials of the electrodes making up the system.

d. Rectangular Hyperbolic Cylinders (see Fig. 11·3*d*).

$$F(w) = Aw^2 + B \qquad [11\cdot30]$$

$$\phi(x,y) = \frac{V_2 - V_1}{C_2^2 - C_1^2}(x^2 - y^2) - \frac{V_2C_1^2 - V_1C_2^2}{C_2^2 - C_1^2} \qquad [11\cdot31]$$

$$\psi(x,y) = 2xy\frac{V_2 - V_1}{C_2^2 - C_1^2} \qquad [11\cdot32]$$

The separations from the origin of the vertices of the two hyperbolic electrodes are C_1 and C_2.

e. Confocal Parabolic Cylinders (see Fig. 11·3*e*).

$$F(w) = Aw^{1/2} + B \qquad [11\cdot33]$$

$$\phi(r,\theta) = \frac{V_2 - V_1}{k_2 - k_1}\left[\frac{r}{2}(1 + \cos\theta)\right]^{1/2} + \frac{V_1k_2 - V_2k_1}{k_2 - k_1} \qquad [11\cdot34a]$$

$$\psi(r,\theta) = \frac{V_2 - V_1}{k_2 - k_1}\left[\frac{r}{2}(1 - \cos\theta)\right]^{1/2} \qquad [11\cdot34b]$$

The focal lengths of the two parabolas corresponding to the electrodes are k_1^2 and k_2^2.

f. Tangent Circular Cylinders (see Fig. 11·3*f*).

$$F(w) = \frac{A}{w + B} \qquad [11\cdot35]$$

$$\phi(r,\theta) = 2R_2R_1\frac{V_1 - V_2}{R_2 - R_1}\frac{\cos\theta}{r} + \frac{V_2R_2 - V_1R_1}{R_2 - R_1} \qquad [11\cdot36]$$

$$\psi(r,\theta) = -2R_2R_1\frac{V_2 - V_1}{R_2 - R_1}\frac{\sin\theta}{r} \qquad [11\cdot37]$$

The radii of the cylindrical electrodes are R_2 and R_1.

g. Slit (see Fig. 11·3*g*).

$$F(w) = A(B + w^2)^{\frac{1}{2}} + Cw + k \qquad [11\cdot38]$$

$$\phi(x,y) = \frac{1}{\sqrt{8}}(E_1 - E_2)\left\{\left[\left\{\left(\frac{D}{2}\right)^2 + x^2 - y^2\right\}^2 + 4x^2y^2\right]^{\frac{1}{2}} + \left(\frac{D}{2}\right)^2 + x^2 - y^2\right\}^{\frac{1}{2}} - \frac{1}{2}(E_1 + E_2)x + V_o \qquad [11\cdot39]$$

$$\psi(x,y) = \frac{1}{\sqrt{8}}(E_1 - E_2)\left\{\left[\left\{\left(\frac{D}{2}\right)^2 + x^2 - y^2\right\}^2 + 4x^2y^2\right]^{\frac{1}{2}} - \left(\frac{D}{2}\right)^2 - x^2 + y^2\right\}^{\frac{1}{2}} - \frac{1}{2}(E_1 + E_2)y \qquad [11\cdot40]$$

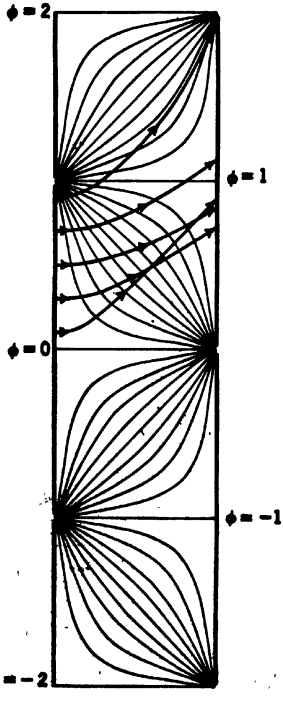


FIG. 11·4. Potential Variation and Electron Paths between Staggered Plane Electrodes.

Here D is the width of the slit and V_o its potential. The fields on either side of the slit at a considerable distance from it (very large positive and negative values of x) are assumed to be uniform and have values of E_2 and E_1, respectively.

There are, of course, a large number of other complex functions which are useful for this process of conformal mapping. Additional examples will be found in textbooks on electricity and potential fields.[1]

11·4. Problems Involving Separation of the Variables. The determination of the potential distribution by means of the separation of the variables, though frequently laborious, particularly when it gives rise to integrals which must be evaluated by numerical integration, can usually be carried out in a fairly straightforward manner. It does not require choosing some special function which fits or closely approximates the conditions of the problem as does the method of conformal representation. However, occasionally, when integrals

[1] See, e.g., Jeans, reference 1, and Ollendorf, reference 2.

obtained during the course of solution cannot be exactly integrated, much loss of time and labor can be avoided by the judicious choice of an analytic function which is sufficiently like the exact function to be substituted for it.

To illustrate the application of this method to a problem encountered in practice, the determination of the potential distribution of the configuration shown in Fig. 11·4 will be made. This configuration corresponds to that which might be used in an elementary form of electrostatic secondary-emission multiplier. The system consists of two rows of plane parallel plates, each of width b, mounted so that the electrodes of the two rows are staggered with respect to each other. The potential increases progressively from plate to plate in uniform steps of V. The separation between the rows is assumed to be d, and for the sake of simplicity the rows are assumed to be indefinitely extended.

It is evident from the configuration that for any constant value of x the potential is periodic in y with the period b, except for the term $2Vy/b$ which represents the average rise in potential with y. Therefore, by using $k = 2\pi n/b$ in Eq. 11·14, the potential may be expressed by

$$\phi(x,y) = \frac{2Vy}{b} + \sum_{n=1}^{\infty}\left(A_n \sinh\frac{2\pi n}{b}x + B_n \cosh\frac{2\pi n}{b}x\right) \cdot \left(\sin\frac{2\pi n}{b}y + D_n \cos\frac{2\pi n}{b}y\right) \qquad [11\cdot41]$$

The coefficients A_n, B_n, and D_n can be determined from the known potential distributions along the two rows of electrodes, at $x = 0$ and $x = d$, by the usual methods of evaluating Fourier coefficients.[2]

For $x = 0$:

$$D_n B_n = \frac{2}{b}\int_{-b/2}^{b/2}\left(\phi(0,y) - \frac{2Vy}{b}\right)\cos\frac{2\pi n}{b}y\,dy = 0 \qquad [11\cdot42]$$

$$B_n = \frac{2}{b}\int_{-b/2}^{b/2}\left(\phi(0,y) - \frac{2Vy}{b}\right)\sin\frac{2\pi n}{b}y\,dy = (-1)^n\frac{4V}{2\pi n} \qquad [11\cdot43]$$

and for $x = d$:

$$A_n \sinh\frac{2\pi nd}{b} + B_n \cosh\frac{2\pi nd}{b} = \frac{2}{b}\int_{-b/2}^{b/2}\left(\phi(d,y) - \frac{2Vy}{b}\right)\sin\frac{2\pi ny}{b}dy = \frac{4V}{2\pi n} \qquad [11\cdot44]$$

The coefficient involving D_n is zero, since D_n itself is zero.

[2] See Adams, reference 3, Formula 6.800.

The final series expressing the potential distribution is obtained by evaluating the coefficients, substituting their values in Eq. 11·41, and simplifying the results with the aid of the addition theorems for hyperbolic functions.[3] The distribution is found to be given by

$$\phi(x,y) =$$

$$\frac{2Vy}{b} + \frac{2V}{\pi}\left\{ - \frac{\sinh \frac{2\pi}{b}\left(\frac{d}{2} - x\right)}{\sinh \frac{\pi d}{b}} \sin \frac{2\pi y}{b} + \frac{1}{2} \frac{\cosh \frac{4\pi}{b}\left(\frac{d}{2} - x\right)}{\cosh \frac{2\pi d}{b}} \sin \frac{4\pi y}{b} \right.$$

$$\left. - \frac{1}{3} \frac{\sinh \frac{6\pi}{b}\left(\frac{d}{2} - x\right)}{\sinh \frac{3\pi d}{b}} \sin \frac{6\pi y}{b} + \frac{1}{4} \frac{\cosh \frac{8\pi}{b}\left(\frac{d}{2} - x\right)}{\cosh \frac{4\pi d}{b}} \sin \frac{8\pi y}{b} \cdots \right\} \qquad [11\cdot45]$$

As has already been pointed out, the method of the separation of variables also finds application to problems where cylindrical polar coordinates are convenient. Where the potential distribution is known on the surface of a circular cylinder, this approach is very effective. Under these conditions Eq. 11·16 representing the distribution reduces to

$$\phi(r,\theta) = \sum_{n=0}^{\infty} \left(\frac{r}{R}\right)^n (a_n \sin n\theta + b_n \cos n\theta) \qquad [11\cdot46]$$

where

$$a_n = \frac{1}{\pi} \int_0^{2\pi} \phi(R,\theta) \sin n\theta \, d\theta \qquad [11\cdot47]$$

$$b_0 = \frac{1}{2\pi} \int_0^{2\pi} \phi(R,\theta) \, d\theta$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} \phi(R,\theta) \cos n\theta \, d\theta, \quad n \neq 0 \qquad [11\cdot48]$$

Here R is the radius of the cylinder on which the potential is known, and $\phi(R,\theta)$ the potential distribution on the mantel surface of this cylinder. This may be illustrated by the simple case of two semicylinders of radius R at potentials V and $-V$, respectively, as shown in Fig. 11·5.

[3] See Adams, reference 3, Formula 3.34.

The coefficients for this configuration become

$$a_n = \frac{1}{\pi}\int_0^{\pi} V \sin n\theta \, d\theta + \frac{1}{\pi}\int_{\pi}^{2\pi} (-V) \sin n\theta \, d\theta = \frac{4V}{\pi n} \ (n \text{ odd}) \qquad [11\cdot49]$$

$$= 0 \ (n \text{ even})$$

$$b_n = \frac{1}{\pi}\int_0^{\pi} V \cos n\theta \, d\theta + \frac{1}{\pi}\int_{\pi}^{2\pi} (-V) \cos n\theta \, d\theta = 0 \qquad [11\cdot50]$$

The series representation, after substituting $n = 2k - 1$, becomes

$$\phi(r,\theta) = \frac{4V}{\pi} \sum_{k=1}^{\infty} \frac{1}{2k-1} \left(\frac{r}{R}\right)^{2k-1} \sin (2k-1)\,\theta \qquad [11\cdot51]$$

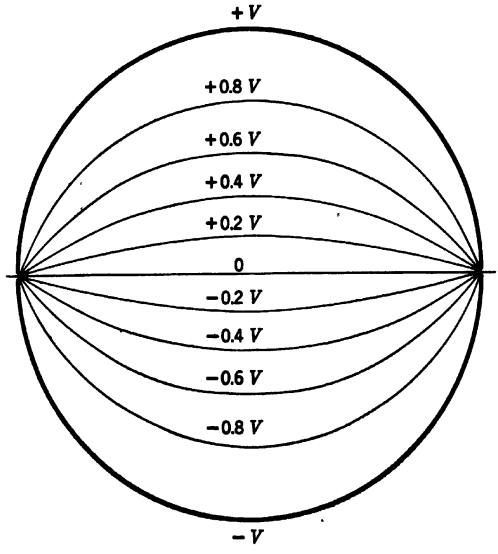


FIG. 11·5. Potential Distribution in Electrode Configuration Consisting of Two Half-Cylinders.

11·5. Special Relationships in the Two-Dimensional Field. Certain relationships and transformations are occasionally helpful in determining the potential distribution in a two-dimensional electrode configuration.

Two-dimensional electron-optical systems which are symmetrical about a plane parallel to the generatrices of the electrodes are very frequently encountered in practice, and this symmetry gives fields within the system rather special properties. A series of parallel slits having a

common axis is an example of this type of electrode arrangement. Other examples are pairs of plates, the plane of symmetry passing midway between them, and a plane and row of rods parallel to it, such as might be used to represent a plane cathode and a control grid (Fig. 11·6).

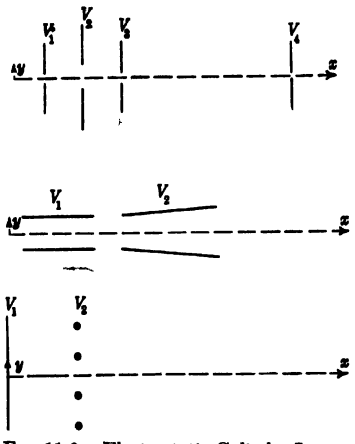


FIG. 11·6. Electrostatic Cylinder Lenses.

To investigate the nature of this type of distribution, let it be assumed that the plane of symmetry is the xz-plane, and that the potential $\phi(x,y)$ is expressed as a power series in y. Since y is the distance from the plane of symmetry all odd-power terms in y must be zero. Therefore the series can be written as

$$\phi(x,y) = \sum_{n=0}^{\infty} A_n(x)y^{2n} \quad [11\cdot52]$$

If this is substituted in the Laplace differential equation (Eq. 11·4), it yields the following relation:

$$\sum_{n=0}^{\infty} \{A_n''(x)y^{2n} + 2n(2n-1)A_n(x)y^{2n-2}\} = 0 \quad [11\cdot53]$$

It can easily be seen, by grouping like powers of y, that the coefficients are related by the following recurrence relations:

$$A_{n-1}''(x) = -2n(2n-1)A_n(x) \qquad n = 1, 2, 3 \cdots \quad [11\cdot54]$$

If the potential distribution along the axis, that is, that for $y = 0$, is $\Phi(x)$, it is evident from Eq. 11·52 that

$$A_0(x) = \Phi(x)$$

By applying the recurrence relation, the other coefficients are found to be:

$$A_1(x) = -\frac{A_0''(x)}{2\cdot1} = -\frac{1}{2}\Phi''(x)$$

$$A_2(x) = -\frac{A_1''(x)}{4\cdot3} = \frac{1}{4!}\Phi^{IV}(x)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$A_n(x) = (-1)^n \frac{1}{(2n)!}\Phi^{(2n)}(x)$$

Therefore Eq. 11·52 can be written

$$\phi(x,y) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} \Phi^{(2n)}(x) y^{2n}$$

$$= \Phi(x) - \frac{1}{2}\Phi''(x)y^2 + \frac{1}{24}\Phi^{IV}(x)y^4 - \cdots \qquad [11\cdot55]$$

This equation makes it possible to determine the entire potential distribution from a knowledge of the distribution along the axis.

Frequently problems involving cylindrical symmetry can be conveniently solved by determining the potential distribution of an analogous plane configuration, and then transforming the solution to fit the desired cylindrical configuration.

If the solution of the plane configuration is given by the function $\phi(x,y)$, then that for the cylindrical structure will be $F(r,\theta) = \phi(\log(r/r_o),\theta)$, the point $x = \log(r/r_o)$, $y = \theta$ in the plane structure corresponding to the point r,θ in the cylindrical structure. Here r_o may be chosen arbitrarily.

This can be seen from an examination of the function $F(r,\theta) = \phi(\log(r/r_o),\theta)$ when substituted in the Laplace equation:

$$\frac{\partial^2\phi\left(\log\frac{r}{r_o},\theta\right)}{\partial\left(\log\frac{r}{r_o}\right)^2} + \frac{\partial^2\phi\left(\log\frac{r}{r_o},\theta\right)}{\partial\theta^2} = r\frac{\partial}{\partial r}r\frac{\partial F}{\partial r} + \frac{\partial^2 F}{\partial\theta^2} \qquad [11\cdot56]$$

Therefore, if $\phi(x,y)$ satisfies the Laplace equation in the form of Eq. 11·4, then $F(r,\theta)$ will satisfy it in the form of Eq. 11·5.

Figure 11·7 illustrates the application of this transformation to the determination of the potential distribution in a special cylindrical triode from that of a plane triode. Here the cathode is represented by the plane mm' and the grid and plate by the set of rods kk' and the plane ll', respectively. When transformed, the elements become a series of concentric electrodes, as shown in the figure. Obviously, the reverse transformation, that is, from a cylindrical system to a plane system, may also be made.

Finally, it is frequently useful to be able to predict the manner in which two equipotentials will intersect. This is particularly true where the equipotentials are obtained by electrolytic plotting, which will be discussed in section 11·10, because the fields at such an intersection vanish so that the equipotentials are hard to locate.

If two equipotentials are assumed to intersect at the point x_o,y_o and

the potential $\phi(x,y)$ is expanded in a Taylor's series about this point,

$$\phi(x,y) = \phi(x_o,y_o) + (x - x_o)\frac{\partial\phi}{\partial x}(x_o,y_o) + (y - y_o)\frac{\partial\phi}{\partial y}(x_o,y_o) + \tfrac{1}{2}(x - x_o)^2\frac{\partial^2\phi}{\partial x^2}(x_o,y_o) + (x - x_o)(y - y_o)\frac{\partial^2\phi}{\partial x\partial y}(x_o,y_o) + \tfrac{1}{2}(y - y_o)^2\frac{\partial^2\phi}{\partial y^2}(x_o,y_o) + \cdots \quad [11\cdot57]$$

the possible angles of intersection of the equipotentials may be determined. Along these equipotentials, obviously, $\phi(x,y) = \phi(x_o,y_o)$ and,

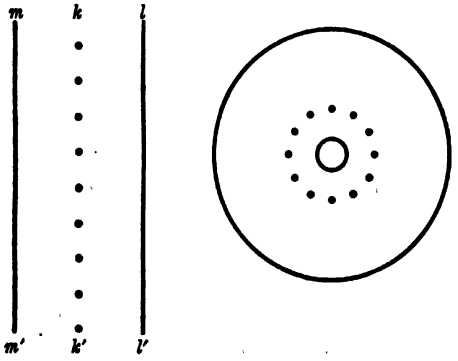


FIG. 11·7. Corresponding Structures of a Plane and a Cylindrical Triode.

furthermore, at x_o,y_o, the first derivatives must be zero. Therefore the equipotentials are the surfaces given by the solutions of the equation

$$(x - x_o)^2\frac{\partial^2\phi}{\partial x^2}(x_o,y_o) + 2(x - x_o)(y - y_o)\frac{\partial^2\phi}{\partial x\partial y}(x_o,y_o) + (y - y_o)^2\frac{\partial^2\phi}{\partial y^2}(x_o,y_o) = 0 \quad [11\cdot58]$$

By transforming the variables to

$$x - x_o = \zeta \qquad y - y_o = \xi \quad [11\cdot59]$$

and putting

$$\frac{\partial^2\phi}{\partial x^2}(x_o,y_o) = a \qquad \frac{\partial^2\phi}{\partial x\partial y}(x_o,y_o) = b \qquad \frac{\partial^2\phi}{\partial y^2}(x_o,y_o) = c \quad [11\cdot60]$$

Eq. 11·58 may be written

$$\zeta^2 a + 2\zeta\xi b + \xi^2 c = 0 \qquad [11\cdot61]$$

When solved, this equation has two real roots corresponding to the two intersecting equipotentials. Since the product of the slopes of the two corresponding lines is

$$\frac{a}{c} = \left[\frac{\dfrac{\partial^2\phi}{\partial x^2}}{\dfrac{\partial^2\phi}{\partial y^2}}\right]_{(x_0,y_0)} = -1 \qquad [11\cdot62]$$

as follows from the fulfilment of the Laplace equation, the two equipotentials must intersect at 90 degrees for a two-dimensional system of the type described (Fig. 11·8).

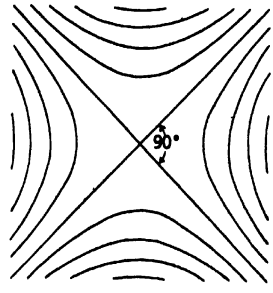


FIG. 11·8. Intersection of Two Equipotentials in a Two-Dimensional Field.

It is not possible for any odd number of equipotentials to come together at a point, as this would require a discontinuity in the derivatives of the potential field, making it impossible to satisfy the Laplace equation. However, any even number may meet, the point of junction being the intersection of continuous equipotential lines. In terms of a Taylor expansion, such as Eq. 11·57, the intersection of three equipotential lines occurs at a higher-order singularity where the first and second derivatives vanish. The intersection of four equipotential lines occurs at a point where the first, second, and third derivatives vanish. Such higher-order singularities occur rarely in practice.

11·6. Axially Symmetrical Systems. By far the most important potential distributions from the standpoint of electron optics involve axial symmetry. Again, no general solution can be given for cases with arbitrary boundary conditions. Indeed, as with two-dimensional systems, frequently no analytic solution is possible.

The Laplace equation in the cylindrical coordinates z,r,θ takes the following form:

$$\frac{\partial^2\phi}{\partial z^2} + \frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial\phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2} = 0 \qquad [11\cdot63]$$

If the z-axis is an axis of symmetry, the derivatives with respect to the angular coordinate must vanish. The differential equation therefore

becomes

$$\frac{\partial^2\phi}{\partial z^2} + \frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial\phi}{\partial r} = 0 \qquad [11\cdot64]$$

The general solution of this equation has the form[4]

$$\phi(z,r) = \frac{1}{2\pi}\int_0^{2\pi} \Phi(z + ir\sin\alpha)d\alpha \qquad [11\cdot65]$$

Again a series of solutions of the Laplace equation can be obtained by substituting simple functions for $\Phi(z)$, the potential distribution along the axis of symmetry. However, the most useful analytical approach for axially-symmetric systems employs the separation of variables. The solution is expressed as the product of functions of z and r as follows:

$$\phi(z,r) = F(z)\cdot G(r) \qquad [11\cdot66]$$

If this is substituted into Eq. 11·64 two differential equations are obtained:

$$\frac{d^2F}{dz^2} = -k^2F \qquad [11\cdot67]$$

$$\frac{1}{r}\frac{d}{dr} r \frac{dG}{dr} = k^2G \qquad [11\cdot68]$$

related to each other by the separation parameter k.

The solution of these two equations may be written in the following form:

$$F(z) = A\sin kz + B\cos kz \qquad [11\cdot69]$$

$$G(r) = CJ_0(ikr) + DN_0(ikr) \qquad [11\cdot70]$$

where $J_0(r)$ is the Bessel function of zero order, which has the well-known form[5]

$$J_0(r) = \sum_{n=0}^{\infty} \frac{1}{(n!)^2}\left(\frac{ir}{2}\right)^{2n} \qquad [11\cdot71]$$

and $N_0(r)$ is the zero-order Neumann function. The latter may be expressed in terms of Bessel functions with the aid of the following relation:[5]

$$N_0(r) = \frac{2}{\pi}\left\{\left(\log\frac{r}{2} + \gamma\right)J_0(r) - 4\sum_{n=1}^{\infty}(-1)^n\frac{1}{2n}J_{2n}(r)\right\} \qquad [11\cdot72]$$

[4] See Brüche and Scherzer, reference 4, p. 65. Geometrische Elektronenoptik
[5] See Jahnke and Emde, reference 5, pp. 196–198.

where

$$\gamma = 0.5772 \quad \text{and} \quad J_n(r) = \left(\frac{r}{2}\right)^n \sum_{k=0}^{\infty} \frac{1}{k!(n+k)!}\left(\frac{ir}{2}\right)^{2k} \qquad [11\cdot73]$$

The general solution is obtained by summing the product of these two functions over all possible values of k both real and imaginary, and thus may be written as follows:

$$\phi(z,r) = \sum_k [A_k \sin kz + B_k \cos kz]\,[J_0(ikr) + D_k N_0(ikr)] \qquad [11\cdot74]$$

The coefficients are evaluated by the usual method from the boundary conditions. Frequently the conditions of the problem are such that the values of k are continuous rather than discrete, and the summation giving the potential distribution becomes an integral, as will be illustrated in the next section.

As was found to be the case for symmetrical two-dimensional systems, the existence of a high degree of symmetry imposes certain special properties upon the potential distribution. If a power series in r is substituted in Eq. 11·64, as was done in deriving Eq. 11·55, an expression describing the entire potential distribution in terms of the potential along the axis, $\phi(z)$, is obtained. This expression has the following form:

$$\phi(z,r) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \Phi^{(2n)}(z) \left(\frac{r}{2}\right)^{2n}$$

$$= \Phi(z) - \frac{r^2}{4}\Phi''(z) + \frac{r^4}{64}\Phi^{IV}(z) \cdots \qquad [11\cdot75]$$

If this is compared with Eq. 11·55 it will be seen that the potential along the axis of axially symmetrical systems follows the potential of the electrodes more closely than the axial potential for two-dimensional systems. This might be expected from the geometry of the two configurations.

The manner in which equipotentials intersect at a saddle point on the axis can be determined from Eq. 11·75. If a Taylor expansion of the potential on the axis,

$$\Phi(z) = \Phi(z_o) + \frac{(z - z_o)^2}{2}\Phi''(z_o) \cdots \qquad [11\cdot76]$$

is substituted in Eq. 11·75, the equipotential surfaces near the axis are given by

$$\frac{(z - z_o)^2}{2} - \frac{r^2}{4} = 0 \qquad r = \pm 2^{\frac{1}{2}}(z - z_o) \qquad [11\cdot77]$$

The equipotentials, therefore, consist of a double cone having a half angle α at the vertex:

$$\alpha = \arctan 2^{1/2} = 54° 44' \qquad [11·78]$$

This type of saddle point is illustrated in Fig. 11·9.

Off the axis equipotential surfaces intersect at right angles, being simultaneously perpendicular to the meridional plane. Their behavior here is thus the same as in two-dimensional systems.

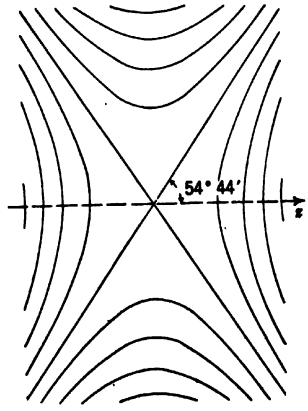


FIG. 11·9. Saddle Point on the Axis of an Axially Symmetric System.

11·7. Typical Configurations with Axial Symmetry. A frequently occurring configuration consists of two semi-infinite coaxial cylinders of equal radius R at potentials V_2 and V_1, as shown in Fig. 11·10*a*. From a consideration of this figure, it will be apparent that the electrode system may be considered as being made up of two separate parts, each consisting of a semi-infinite cylinder terminated by a plane whose potential is $(V_1 + V_2)/2$. The two cylinders are at potentials V_1 and V_2, and the common terminating plane may conveniently be placed at the origin $z = 0$ of the coordinate system.

The problem of a semi-infinite cylinder at a given potential terminated in a plane at a second potential is a special case of the more general problem wherein the cylinder may have any potential distribution $V(z)$ along its mantel surface, where $V(z)$ is, of course, assumed to be finite throughout. If V_o is the potential of the terminating disk, the

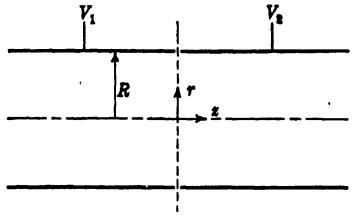


FIG. 11·10(a). Electrode Configuration Consisting of Coaxial Equidiameter Cylinders.

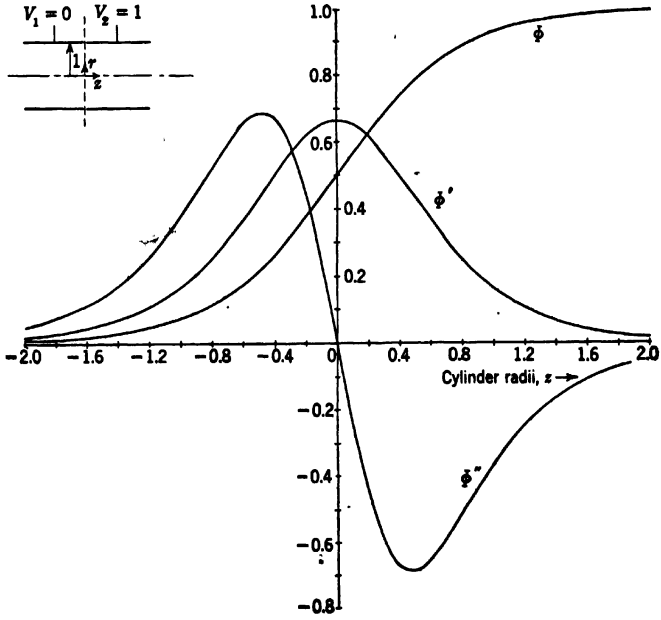


FIG. 11·10(b). Variation of the Potential on the Axis and Its First Two Derivatives in the Electron Lens Formed at the Junction of Two Coaxial Equidiameter Cylinders.

general solution becomes

$$\phi(z,r) = V_o + \int [A_k \sin kz + B_k \cos kz][J_0(ikr) + D_k N_0(ikr)]dk \qquad [11\cdot79]$$

Since $\phi(z,r) = V_o$ at $z = 0$, B_k must be equal to zero. Furthermore, since the Neumann function $N_0(ikr)$ becomes infinite when $r = 0$, the coefficient D_n must vanish. Finally, since ϕ is finite for very large values of k, the separation parameter must be real. Therefore the general solution reduces to

$$\phi(z,r) = V_o + \int_0^\infty A_k \sin kz\, J_0(ikr)dk \qquad [11\cdot80]$$

The coefficient A_k can be evaluated in the usual way by the integral

$$A_k = \frac{1}{J_0(ikR)} \frac{2}{\pi} \int_0^\infty (V(z) - V_o) \sin kz\, dz \qquad [11\cdot81]$$

For the particular problem where $V_o = (V_1 + V_2)/2$ and the potential at $r = R$ is $V(z) = V_2$ for positive z and $V(z) = V_1$ for negative z, the coefficient becomes[6]

$$A_k = \frac{1}{J_0(ikR)} \frac{2}{\pi} \int_0^\infty \left(V_2 - \frac{V_2 + V_1}{2}\right) \sin kz\, dz$$

$$= \frac{1}{J_0(ikR)} \frac{2}{\pi k} \frac{V_2 - V_1}{2} \qquad [11\cdot82]$$

Hence the solution is

$$\phi(z,r) = \frac{V_2 + V_1}{2} + \frac{V_2 - V_1}{\pi} \int_0^\infty \frac{\sin kz}{k} \frac{J_0(ikr)}{J_0(ikR)} dk \qquad [11\cdot83a]$$

On the axis the distribution becomes

$$\Phi(z) = \frac{V_2 + V_1}{2} + \frac{V_2 - V_1}{\pi} \int_0^\infty \frac{\sin kz}{k} \frac{dk}{J_0(ikR)} \qquad [11\cdot83b]$$

This distribution is shown in Fig. 11·10*b* together with its first two derivatives. The functions were evaluated by numerical integration with R taken as the unit of length.

A second example to be considered is the field of an electrode system

[6] The term of A_k arising from substituting $z \to \infty$ in the integrated expression does not contribute to the potential as given by Eq. 11·80 since it represents an infinitely rapid oscillation with k.

used in an image tube described in section 2·2. The configuration is shown in Fig. 11·11. It consists of two coaxial equidiameter cylinders, one of which is terminated by a plane disk at $z = 0$. The potential of this disk is V_o. Adjoining it is the first cylinder of radius R, having a length a along which the potential varies uniformly from $V(z) = V_o$ at $z = 0$ to $V(z) = V_1$ at $z = a$. The second cylinder is assumed to extend from $z = a$ to infinity and to be at the constant potential V_2.

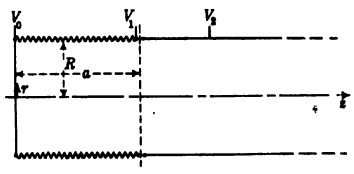


FIG. 11·11. Electrode Configuration of an Electrostatic Image Tube.

Again all coefficients except A_k vanish. This coefficient is given by

$$A_k = \frac{1}{J_0(ikR)} \frac{2}{\pi} \left\{ \int_0^a \frac{V_1 - V_o}{a} z \sin kz\, dz + \int_a^\infty (V_2 - V_o) \sin kz\, dz \right\} \qquad [11·84]$$

$$= \frac{1}{J_0(ikR)} \frac{2}{\pi} \left\{ \frac{V_1 - V_o}{ak^2} \sin ak + \frac{V_2 - V_1}{k} \cos ak \right\}$$

If this is substituted in Eq. 11·80, the potential distribution becomes

$$\phi(z,r) = V_o + \frac{2}{\pi} \int_0^\infty \left\{ \frac{V_1 - V_o}{ak^2} \sin ak + \frac{V_2 - V_1}{k} \cos ak \right\} \frac{J_0(ikr)}{J_0(ikR)} \sin kz\, dk \qquad [11·85]$$

This solution may also be used to estimate the potential distribution for two semi-infinite coaxial equidiameter cylinders separated from each other by a finite small distance d if the assumption is made that at a radius R the potential varies uniformly from V_1, the potential of the left-hand cylinder, to V_2, that of the right-hand cylinder.[7] This electrode system is illustrated in Fig. 11·12. If the origin is placed at the mid-point between the two cylinders, Eq. 11·85

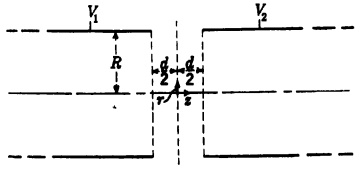


FIG. 11·12. Electrode Configuration for Lens Formed by Two Equidiameter Cylinders Separated by a Distance d.

[7] A more accurate description of the potential variation between the edges of the cylinder requires resorting to the electrolytic plotting tank. See Bertram, reference 6. For a convenient general method for dealing with systems of this type, see Bertram, reference 7.

must apply, provided that here a is replaced by $d/2$, V_o by $(V_1 + V_2)/2$, and V_1 by V_2. Thus

$$\phi(z,r) = \frac{V_1 + V_2}{2} + \frac{2}{\pi}\int_0^\infty \frac{V_2 - V_1}{dk^2} \sin\frac{dk}{2}\frac{J_0(ikr)}{J_0(ikR)} \sin kz\, dk \qquad [11\cdot86]$$

For a region bounded by a cylinder whose radius is R and whose potential is given by $V(z)$, and which is terminated at the ends by disks at potentials V_1 and V_2, a Fourier series must be used instead of the Fourier integrals of the preceding examples. Thus, in the electrode system of this type which is illustrated in Fig. 11·13, where the length of the cylinder is b, the potential distribution becomes

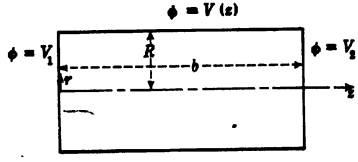


FIG. 11·13. Electrode Configuration for a Cylinder Terminated at Both Ends by Disks of Constant Potential.

$$\phi(z,r) = V_1 + \frac{V_2 - V_1}{b} z + \sum_{n=1}^{\infty} A_n \sin\frac{\pi n z}{b} J_0\left(\frac{i\pi n r}{b}\right) \qquad [11\cdot87a]$$

where

$$A_n = \frac{1}{J_0\left(\dfrac{i\pi n R}{b}\right)} \frac{2}{b}\int_0^b \left\{V(z) - V_1 - \frac{V_2 - V_1}{b} z\right\} \sin\frac{\pi n z}{b}\, dz \qquad [11\cdot87b]$$

Another type of system which is of some interest is one which is nearly the inverse of that described by Eq. 11·83. This system, which is illustrated in Fig. 11·14, consists of a semi-infinite cylinder of radius R at potential V_o terminated at the origin $z = 0$ by a disk over which the potential is given by the function $\phi(0,r) = V(r)$. For this configuration the separation parameter must be imaginary since $\phi - V_o$ becomes zero at $r = R$ and the Bessel function $J_0(ix)$ is greater than zero for all real values of x. Hence the potential becomes

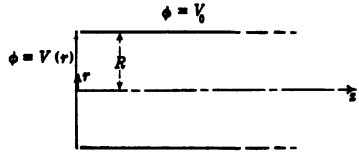


FIG. 11·14. Cylinder at Uniform Potential Terminated by a Disk with Prescribed Potential Distribution.

$$\phi(z,r) = V_o + \sum_{n=1}^{\infty} A_n \exp\left(-\frac{\mu_n z}{R}\right) J_0\left(\frac{\mu_n r}{R}\right) \qquad [11\cdot88a]$$

where μ_n is the nth root of $J_0(x)$ and

$$A_n = \frac{2}{J_1^2(\mu_n)} \frac{1}{R^2} \int_0^R (V(r) - V_o) J_0\left(\frac{\mu_n r}{R}\right) r\, dr \qquad [11\cdot88b]$$

11·8. Selection of Coordinates. The particular solution of the differential equation giving the potential distribution for a specified set of boundary conditions is frequently obtained most readily if an electrode that is at a known potential can be made to coincide with one of the coordinate surfaces of the coordinate system used to express the distribution. Thus, if the methods using cylindrical coordinates, discussed in the two preceding sections, fail, it is helpful to choose another set of coordinates which fulfil this condition.

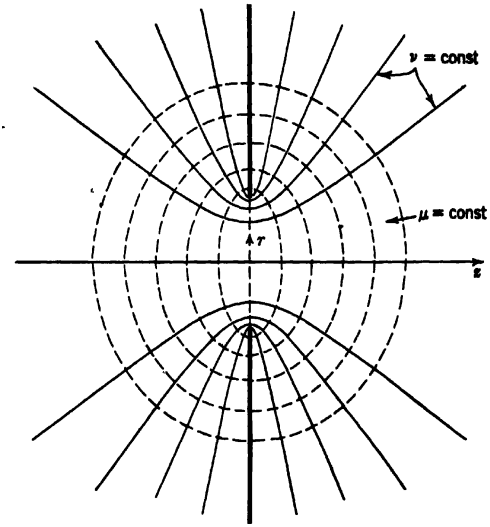


FIG. 11·15. Coordinate System for Determining Potential Distribution about Circular Aperture.

This is illustrated by the example of a single circular aperture of radius R and at potential V_o with a field which approaches E_1 away from the aperture on one side and E_2 on the other.[8] The coordinate system which best fits this problem consists of confocal hyperboloids and conjugate confocal ellipsoids (Fig. 11·15). In this system, if

$$\frac{r^2}{R^2(1 - \nu^2)} - \frac{z^2}{R^2\nu^2} = 1 \qquad [11\cdot89]$$

[8] See Brüche and Scherzer, reference 4, p. 69.

where r and z correspond to the ordinary cylindrical coordinates and ν is the hyperboloid coordinate, the electrode coincides with the surface

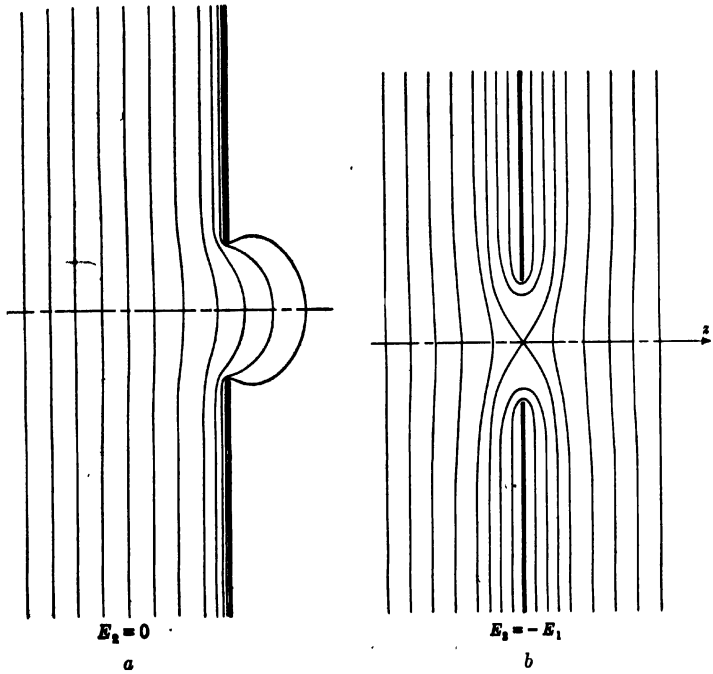


FIG. 11·16. Potential Distribution about Circular Aperture.

$\nu = 0$. The conjugate ellipsoids for this system are the surfaces μ = const of the equation

$$\frac{r^2}{R^2(\mu^2 + 1)} + \frac{z^2}{R^2\mu^2} = 1 \qquad [11\cdot90]$$

The relation between the new coordinates (μ,ν) and the old cylindrical coordinates (r,z) is

$$r = R[(\mu^2 + 1)(1 - \nu^2)]^{\frac{1}{2}}$$
$$\mu^2 \geq 0 \qquad \nu^2 \leq 1 \qquad z = R\mu\nu \qquad [11\cdot91]$$

The Laplace equation in the new coordinates can, hence, be shown to be

$$\frac{\partial}{\partial\mu}\left\{(\mu^2 + 1)\frac{\partial\phi}{\partial\mu}\right\} + \frac{\partial}{\partial\nu}\left\{(1 - \nu^2)\frac{\partial\phi}{\partial\nu}\right\} = 0 \qquad [11\cdot92]$$

Again the method of separating the variables,

$$\phi(\mu,\nu) = M(\mu)\cdot N(\nu)$$

which has been found to be so fruitful in previous problems, can be used. This leads to the general solution

$$\phi(\mu,\nu) = A\mu\nu + B(\mu \arctan \mu + 1) + C \qquad [11\cdot93]$$

If the boundary conditions and retransforming to cylindrical coordinates are taken into account, the potential distribution is found to be

$$\phi(z,r) = V_o - \frac{E_1 + E_2}{2} z + \frac{1}{\pi}(E_1 - E_2)|z|\left(\arctan \mu + \frac{1}{\mu}\right)$$

where [11·94]

$$\mu = \left\{\frac{1}{2}\left(\frac{z^2}{R^2} + \frac{r^2}{R^2} - 1\right) + \frac{1}{2}\left[\frac{4z^2}{R^2} + \left(\frac{z^2}{R^2} + \frac{r^2}{R^2} - 1\right)^2\right]^{\frac{1}{2}}\right\}^{\frac{1}{2}}$$

On the z-axis the distribution is

$$\Phi(z) = V_o - \frac{E_1 + E_2}{2} z + \frac{R}{\pi}(E_1 - E_2)\left(\frac{z}{R}\arctan\frac{z}{R} + 1\right) \qquad [11\cdot95]$$

The potential distribution about an aperture for $E_1 = -E_2$ and for $E_1 \neq 0$, $E_2 = 0$ is illustrated in Figs. 11·16*a* and *b*.

Another instance in which a different system of coordinates is convenient is where the potential is defined over concentric spherical shells. Spherical coordinates are most appropriate for this type of problem. The Laplace equation for the most suitable spherical coordinates, which are illustrated in Fig. 11·17, becomes

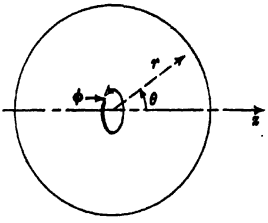


FIG. 11·17. Spherical Coordinates.

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi}{\partial\theta}\right) = 0 \qquad [11\cdot96]$$

The general solution, meeting the requirement that the potential ϕ remain finite on the axis of symmetry (with the possible exception of $r = 0$), is

$$\phi(r,\theta) = \sum_{n=0}^{\infty}\left(A_n r^n + \frac{B_n}{r^{n+1}}\right)P_n(\cos\theta) \qquad [11\cdot97]$$

where $P_n(x)$ is the Legendre polynomial of order n:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \qquad [11\cdot98]$$

The most elementary example of such a spherical system consists of two concentric shells whose radii are r_1 and r_2 at potentials V_1 and V_2, respectively. The potential distribution for this is

$$\phi(r,\theta) = \Phi_o(r) = \frac{V_2 r_2 - V_1 r_1}{r_2 - r_1} - \frac{V_2 - V_1}{r_2 - r_1} \frac{r_1 r_2}{r} \qquad [11\cdot99]$$

Suppose now that by the insertion of ring electrodes the potential over the surface $\theta = \theta_o$ is given some potential variation $V(r)$ instead of the distribution $\Phi_o(r)$ which it would have in the absence of these electrodes. Since the solution $\Phi_o(r)$ satisfies the Laplace equation, and the change in boundary conditions along θ_o can be expanded into the power series

$$P_n (\cos \theta_o)\{V(r) - \Phi_o(r)\} = \sum_{n=0}^{\infty} \left(a_n r^n + \frac{b_n}{r^{n+1}}\right) \qquad [11\cdot100]$$

so as to lead to a solution of the form of Eq. 11·97, the potential distribution becomes

$$\phi(r,\theta) = \Phi_o(r) + \sum_{n=0}^{\infty} \left(a_n r^n + \frac{b_n}{r^{n+1}}\right) \frac{P_n (\cos \theta)}{P_n (\cos \theta_o)} \qquad [11\cdot101]$$

If, on the other hand, the potential distribution within a sphere of radius R is required, where the potential over its surface varies as $V(\theta)$, the solution is found to be

$$\phi(r,\theta) = \sum_{n=0}^{\infty} A_n r^n P_n (\cos \theta) \qquad [11\cdot102a]$$

where

$$A_n = \frac{1}{R^n} \frac{2n+1}{2} \int_0^{\pi} V(\theta) P_n (\cos \theta) \sin \theta \, d\theta \qquad [11\cdot102b]$$

11·9. Graphical Potential Mapping: The Liebmann Procedure. The Liebmann net method provides a simple and relatively speedy procedure for determining the potential distribution in arbitrary two-dimen-

sional electrode systems with prescribed boundary potentials. It involves a process of successive approximations so that the accuracy depends on the skill and judgment exercised, and the number of successive applications of the procedure to the configuration. The Liebmann procedure is of particular interest since it makes possible the very effective utilization of business machines operating with punched cards for the solution of potential problems.[9]

Figure 11·18 represents a two-dimensional configuration for which the potential distribution over the boundary curve A is given by $V(x,y)$. A rectangular coordinate net having a mesh width Δ is drawn over the area enclosed by A. If the potentials ϕ_1, ϕ_2, ϕ_3, ϕ_4, and ϕ_0 at the points denoted on the figure by 1, 2, 3, 4, and 0, respectively are considered, it is evident that

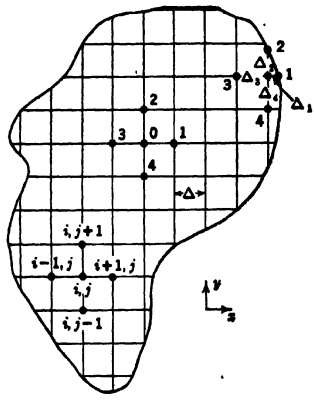


FIG. 11·18. The Liebmann Procedure.

$$\frac{\phi_1 - \phi_0}{\Delta} - \frac{\phi_0 - \phi_3}{\Delta} \cong \left(\frac{\partial \phi}{\partial x}\right)_{10} - \left(\frac{\partial \phi}{\partial x}\right)_{03} \cong \Delta \left(\frac{\partial^2 \phi}{\partial x^2}\right)_0$$

and

$$\frac{\phi_2 - \phi_0}{\Delta} - \frac{\phi_0 - \phi_4}{\Delta} \cong \left(\frac{\partial \phi}{\partial y}\right)_{20} - \left(\frac{\partial \phi}{\partial y}\right)_{04} \cong \Delta \left(\frac{\partial^2 \phi}{\partial y^2}\right)_0$$

if the assumption is made that Δ is small.

Since, from Laplace's equation,

$$\left(\frac{\partial^2 \phi}{\partial x^2}\right)_0 + \left(\frac{\partial^2 \phi}{\partial y^2}\right)_0 = 0$$

it follows that

$$\phi_0 \cong \tfrac{1}{4}(\phi_1 + \phi_2 + \phi_3 + \phi_4) \qquad [11\cdot103]$$

If a more general net point x_i,y_j is considered, the potential $\phi(x_i,y_j)$ will be

$$\phi(x_i,y_j) \cong \tfrac{1}{4}\{\phi(x_{i+1},y_j) + \phi(x_{i-1},y_j) + \phi(x_i,y_{j+1}) + \phi(x_i,y_{j-1})\} \qquad [11\cdot104]$$

[9] See Kormes, reference 8.

Around the edges of the system the separation between the net point whose potential is being determined and the boundary will in general be less than the mesh width Δ. If Δ_1, Δ_2, Δ_3, and Δ_4 are the separations of the four neighboring points from the point considered, the potential at the latter point becomes

$$\phi_0 = \frac{\Delta_1\Delta_2\Delta_3\Delta_4}{\Delta_1\Delta_3 + \Delta_2\Delta_4}\left\{\frac{1}{\Delta_1 + \Delta_3}\left(\frac{\phi_1}{\Delta_1} + \frac{\phi_3}{\Delta_3}\right) + \frac{1}{\Delta_2 + \Delta_4}\left(\frac{\phi_2}{\Delta_2} + \frac{\phi_4}{\Delta_4}\right)\right\} \qquad [11\cdot105]$$

The procedure to be followed in employing the above relations is to assume some reasonable potential distribution $\phi^{(0)}(x_i,y_j)$ over the mesh. Then, by starting from the boundary, a new distribution $\phi^{(1)}(x_i,y_j)$ is determined with the aid of Eqs. 11·104 and 11·105, with the known values on the boundary and the assumed potential being used. After the points adjoining the boundary have been corrected, these improved

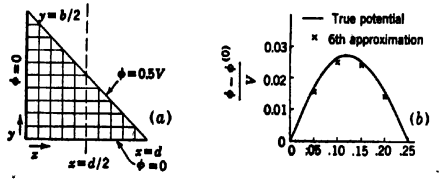


FIG. 11·19. Application of Liebmann Procedure to Specific Electrode Structure. (*a*) Boundaries and Liebmann Net. (*b*) Potential Variation for $x = d/2$.

values are used in determining the next set of points. When the new distribution has been completed, the procedure is repeated, giving a second approximation $\phi^{(2)}(x_i,y_j)$. The process of approximation is continued until the changes in the potential distribution are less than the accuracy required. Further details and refinements in the Liebmann procedure will be found in an article on this subject by Shortley and Weller.[10]

The electrode structure shown in Fig. 11·4 provides a good example for the application of the Liebmann procedure. The symmetry properties of this configuration are such that the potential distribution throughout the system may readily be derived from the potential distribution in the right triangle represented in Fig. 11·19*a*. The relatively wide-meshed Liebmann net employed in the potential deter-

[10] See reference 9.

mination is indicated. It was found that, starting with plausible linear variations of the potential along the vertical and horizontal net lines, six successive approximations sufficed to reduce the deviation from the actual potential along the line $x = d/2$ to less than 1 per cent. This is shown in Fig. 11·19*b*. The actual potential variation may be obtained from Eq. 11·45 by setting $d = b/2$ and $x = d/2$:

$$\phi\left(\frac{d}{2}, y\right) = \frac{2Vy}{b} + \frac{2V}{\pi}\left\{\frac{1}{2}\operatorname{sech}\pi \sin\frac{4\pi y}{b} + \frac{1}{4}\operatorname{sech} 2\pi \sin\frac{8\pi y}{b} + \cdots\right\} \quad [11\cdot106]$$

The initial, assumed, distribution is simply

$$\phi^{(0)}\left(\frac{d}{2}, y\right) = \frac{2Vy}{b} \quad [11\cdot107]$$

11·10. Potential Mapping by Means of the Electrolytic Plotting Tank. The analytical determination of the potential distribution for a given electrode system is always difficult, and frequently impossible. Although a distribution can always be obtained with the aid of the Liebmann procedure, it requires a great deal of time and labor. Fortunately there is an experimental method of making the determination which yields results having an accuracy adequate for a majority of practical problems and is yet fairly easy to carry out.

The method consists of immersing an enlarged model of the electrode system in an electrolyte, applying potentials to these model electrodes which are proportional to those actually used, and measuring, with a suitably constructed probe, the voltage distribution within the model system. The potential distribution measured in this way is that of the actual electrode configuration, as can readily be seen from the following reasoning. When the potential is applied to the electrodes, a current flows through the electrolyte. Obviously, as much current must leave any volume element of the electrolyte as arrives, that is, the current density j must obey the equation of continuity:

$$\frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z} = 0 \quad [11\cdot108]$$

Since the conduction of an electrolyte is ohmic, the current density at any point is proportional to the electric field E:

$$j_x = \sigma E_x \qquad j_y = \sigma E_y \qquad j_z = \sigma E_z \quad [11\cdot109]$$

where σ, the conductivity, is a constant. Hence Eq. 11·108 is equivalent to

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0 \quad [11\cdot110]$$

or, in terms of the potential distribution,

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = 0$$

Therefore the potential distribution within the model obeys the Laplace equation, and, since the boundary conditions are the same as those for the actual electrode system, the measured distribution is the one required.

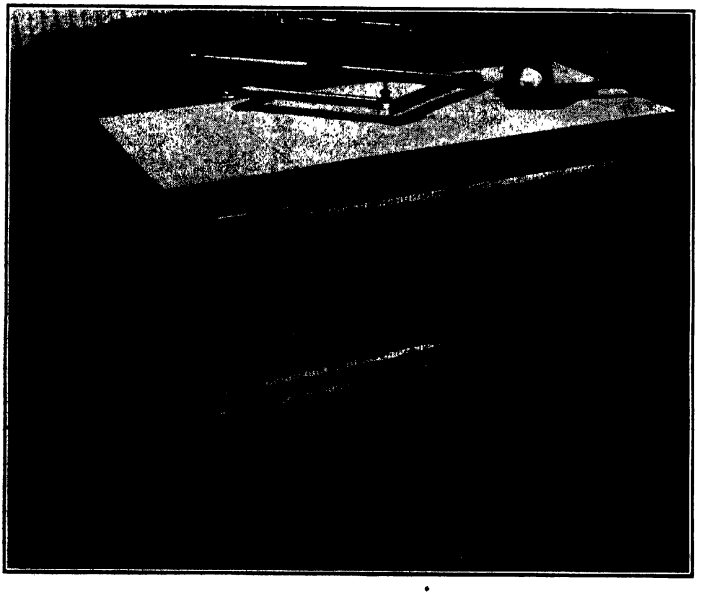


FIG. 11·20. Electrolytic Plotting Tank.

Since, as has already been mentioned, most of the electrode configurations of practical electron-optical systems are either " two-dimensional " or axially symmetric, they have a plane of mirror symmetry. At the plane of symmetry the electric field can have components lying only in this plane, and the equipotentials must intersect it at right angles. Therefore, in the electrolytic model, the current flow across this boundary will be zero, and, if the surface of the electrolyte coincides with the plane of symmetry, the absence of the portion of the model above the liquid does not cause the potential distribution to depart from that in the actual electrode system.

Advantage is taken of this in most practical electrolytic tanks. They

are constructed in such a manner that models of half the electrode system, divided at a plane of symmetry, are immersed so that this plane just coincides with the surface of the electrolyte. The measuring probe moves in the plane of the surface and just touches the liquid.

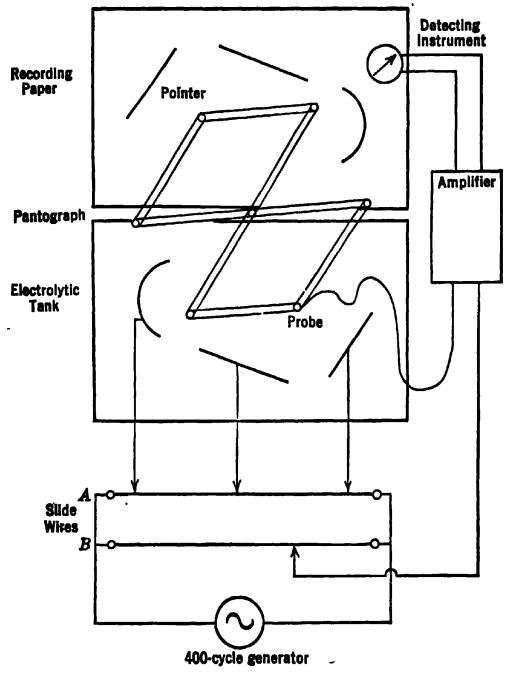


FIG. 11·21. Diagram of Electrolytic Plotting Tank.

Figure 11·20 is a photograph of an electrolytic tank, and Fig. 11·21 is a diagram showing it in plan. The tank itself is 2 feet deep and has a surface area of 2½ by 5½ feet. Although constructed of wood, it is lined with sheet copper, both for purposes of shielding and to insure that it will be water-tight. The lining may be coated with an insulating material such as roofing cement, so that the copper surface is not in contact with the electrolyte, or it may be left exposed. If the tank is fairly large, that is, three or more times the width of the electrode system to be tested, this makes relatively little difference. If the tank is small, the boundary will influence the field distribution about the electrodes, since, if the walls are conducting, they must of necessity be equipotential,

whereas, if they are made of insulating materials, the equipotentials will meet them at right angles to the surface. Usually the latter condition is preferable.

The electrode models may be made of light sheet metal cut and bent to conform with the actual system. They can be supported from above by cross members attached to the walls of the tank. The principal consideration in arranging the supports is not to interfere with the motion of the measuring probe.

In the case of axially symmetric systems the preparation of the electrode models may prove quite time-consuming. The preparation may be greatly simplified by the employment of a *wedge tank*,[11] which utilizes to the fullest degree the symmetry properties of such systems. The wedge tank is a tank with a sloping, insulated bottom filled with electrolyte up to a line on the bottom which corresponds to the axis of the system. The electrolyte thus forms a liquid wedge bounded by the bottom of the tank and by the surface of the electrolyte, both of which correspond to planes of symmetry of the system. If the angle of the wedge is small a sufficiently narrow portion of the circumference of the electrodes is required for the model that the model electrodes may be in the form of flat metal strips. In order to obtain accurate results with this tank the contact angle between the liquid and the tank bottom should approximate the angle of the wedge. This is attained by permitting the liquid to wet the bottom only up to the line representing the system axis and then gradually reducing the wedge angle. It is desirable here to use a large-scale model, so that it does not become necessary to plot the field very near the axis where the capillary rise of electrolyte on the probe may cause serious error.

The probe, as can be seen from the diagram (Fig. 11·21), is carried on a pantograph attached to the side of the tank. This pantograph transmits the motion of the probe to the plotting stylus or pencil which moves over a drawing board. A fine insulated wire runs through the center of the probe and makes contact with the surface of the electrolyte. The barrel of the probe is a metal tube, strong enough to hold the wire rigid and grounded to shield the internal conductor. A shielded flexible cable connects the central conductor with the null-reading instrument.

Potentials are supplied to the various electrodes from slide wires located at any convenient point. The voltage for the probe is obtained from a similar slide wire located within easy reach of the operator from the plotting position, as this must be varied throughout the course of the mapping. The slide wires are clearly visible just below the plotting table in Fig. 11·20.

[11] See Bowman-Manifold and Nicoll, reference 10.

In the plotting tank illustrated, a 400-cycle alternating voltage is applied to the probe and the electrodes in place of a constant voltage. This prevents errors resulting from polarization or the formation of gas bubbles at the electrodes and the measuring probe. It has the added advantage of greatly simplifying the null-reading equipment. The

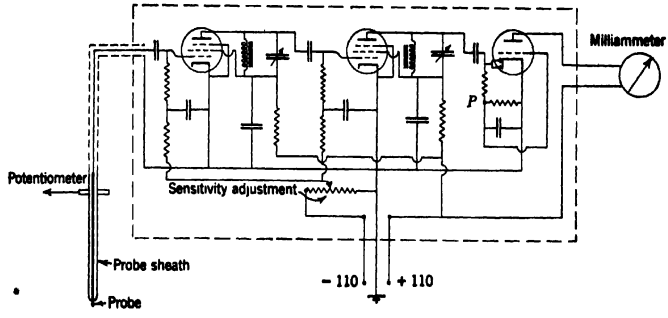


FIG. 11·22. Circuit of Null-Reading Instrument.

frequency is chosen in such a way that there is a minimum of interference between the power lines and the measuring instrument. Ease of generation and measurement is another factor entering into the selection of frequency. It should be mentioned that care must be taken to avoid any phase shift between the voltages supplied to the electrodes and the probe.

The circuit employed in the null-reading device is shown in Fig. 11·22. It consists of an electrostatically shielded tuned amplifier, a compound rectifier and triode (6SR7), and a direct-current milliammeter. In case of unbalance, current is collected by the probe, causing an alternating input voltage to be impressed on the amplifier. The amplified signal is rectified by the diode in the 6SR7 tube, causing a voltage negative with respect to the cathode to appear at the point P, which is connected to the control grid of the triode section of the same tube. If the signal is large enough it will completely suppress the plate current of the triode and, hence, that in the milliammeter. If the probe collects no current whatever, that is, if the voltage applied to the probe sheath is equal to the space potential at the point of immersion of the probe, the voltage at P is just equal to that of the cathode and the plate current of the triode is a maximum.

Thus the null position is indicated as a maximum reading on the meter. The power is supplied by a stable 220-volt direct-current line grounded at the midpoint. The negative half of the supply serves to

adjust the bias on the two (variable-mu) amplifier tubes, permitting an increase in the sensitivity of the device as the point of balance is approached.

The usual procedure in plotting the potential distribution of a given electrode system, after mounting the model and applying suitable potentials to the elements, is this. The overall voltage is subdivided into an appropriate number of equal intervals, the number being dependent upon the accuracy required. The probe potential is then set equal to that corresponding to the first step, and the probe is moved so as to trace

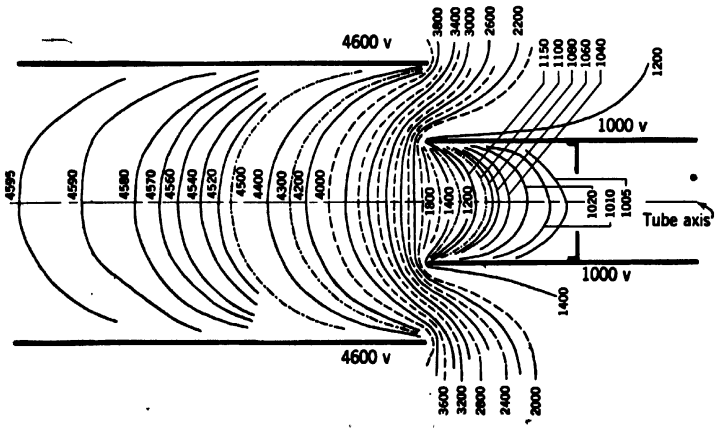


FIG. 11·23. Equipotential Map of the Second Lens of an Electron Gun.

out the equipotential for this voltage on the plotting paper. This procedure is repeated for all the steps. The resulting map shows the equipotential distribution throughout the electrode system. Some representative field maps, obtained in this manner, are shown in Figs. 11·23 and 11·24. The first shows the field distribution in the second lens of an electron gun; the second that in an equidiameter cylinder lens such as is employed in electrostatic image tubes. A map of this type can be used directly for electron trajectory tracing, as will be explained in the next chapter, or may be used to obtain data relating to the potential distribution in any other form required.

For determining the electron-optical properties of axially symmetric systems the axial distribution, together with its first and second derivatives, is necessary. The axial distribution may be obtained directly from a map of the type described. The first derivative may be obtained with a fair degree of accuracy by measuring the slope of the curve express-

ing the axial potential as function of the distance z along the axis. However, the second derivative cannot be derived from the first obtained in this manner without introducing serious error. A consideration of the Laplace equation for axially symmetric systems (Eq. 11·5) shows that the following relation exists:

$$\Phi'' = \frac{2\Phi'}{\rho} \qquad [11\cdot111]$$

where ρ is the radius of curvature of the equipotential at the point in question. Since the radius of curvature of the equipotential can be measured on the map, and the potential gradient Φ' is easily determined, this relation affords at times a means of obtaining a good estimate of the second derivative Φ'' of the potential on the axis.

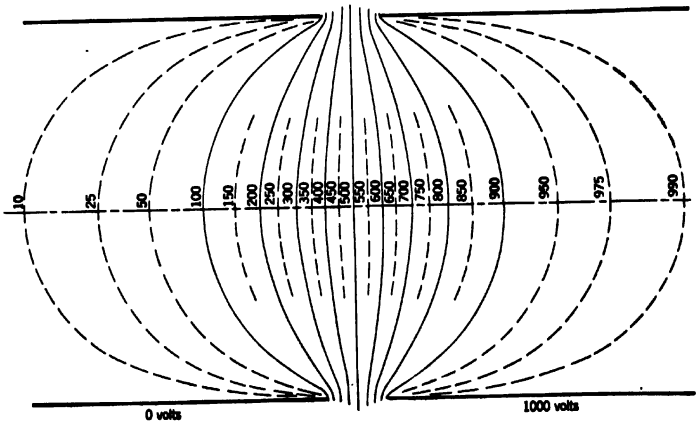


FIG. 11·24. Equipotential Map of an Equidiameter Cylinder Lens.

A more accurate method of accomplishing the same purpose has been suggested by Gundert.[12] Measurements of the potential are made at a series of points in the plane $z = z_o$, normal to the axis. These are plotted with respect to r^2, r being the distance from the axis. Since the potential in the plane in question is

$$\phi(z_o,r) = \Phi(z_o) - \frac{r^2}{4}\Phi''(z_o) + \frac{r^4}{64}\Phi^{IV}(z_o) - \cdots$$

[12] See reference 11.

the slope of the resulting curve at $r = 0$:

$$\frac{d\phi}{d(r^2)} = -\frac{1}{4}\Phi''(z_o) + \frac{r^2}{32}\Phi^{IV}(z_o) - \cdots$$

$$= -\frac{1}{4}\Phi''(z_o) \qquad [11\cdot112]$$

gives, directly, the second derivative on the axis.

11·11. Three-Dimensional Potential Distributions. As has already been pointed out, electron-optical problems requiring the solution of the Laplace equation for the general three-dimensional case are rarely encountered. However, when such problems do occur, it is possible, with certain rather obvious modifications of the electrolytic plotting tank, to determine the potential distribution as outlined in the preceding section. In this case a complete model of the electrode system must be used. The model is submerged below the surface of the electrolyte, the supports being constructed of a dielectric and placed in such a way as to influence the potential distribution as little as possible. The exploring probe must be well insulated and waterproofed, except for a small point which is left exposed to act as the measuring point. The linkage carrying the probe must necessarily be much more complicated than the simple pantograph described in section 11·10.

11·12. Potential Distribution in the Presence of Space Charge. The potential at any point in the neighborhood of a charge q is

$$\phi = \frac{q}{r} \qquad [11\cdot113]$$

where r is the separation between the point where the potential is being determined and the charge. If the charge is distributed throughout a system, the potential is given as the sum of the potentials produced by each element of charge, or, in integral form:

$$\phi = \int \frac{\rho d\tau}{r} \qquad [11\cdot114]$$

where ρ is the charge density, $d\tau$ an element of volume, and r the distance between the point in question and the volume element.

This formulation requires a knowledge of the charge distribution throughout the entire system. Actually, this information is rarely available in practical problems. In such problems, in general, the potentials of the boundaries of the region under consideration are known and, in addition, the amount of free charge within this space. To determine the potential distribution from this information, the usual pro-

cedure is first to calculate the distribution in the absence of space charge, then, with the aid of this potential distribution and the known amount of charge, to calculate the space-charge distribution. Next, with the bounding conductors treated as though they were at zero potential, the charges induced on them by the space charge are determined, and a potential distribution resulting from the charge and its electrical images is calculated. This distribution is added to that obtained in the absence of space charge, yielding the required potential distribution. If the space charge is large, it may be necessary to recalculate the charge distribution, and again to determine the potential distribution to be added. The procedure must be repeated until the charge and potential distributions are self-consistent.

The method of images can be used quite generally in the calculation of the potential distributions resulting from charges in the presence of conductors. The case is simplest if the conductors are plane. If, for example, a charge q is located at a distance d from a conducting plane which can be taken as the yz-plane, the potential distribution in front of the plane is exactly the same as though a charge $-q$ at a distance $-d$ from the plane were present, and the conducting plane were not there (Fig. 11·25). The two charges cooperate to give a potential which vanishes at the yz-plane, given by the following expression:

$$\phi(x,y,z) = q\{([x-d]^2 + y^2 + z^2)^{-\frac{1}{2}} - ([x+d]^2 + y^2 + z^2)^{-\frac{1}{2}}\} \qquad [11\cdot115]$$

If the charge is between two parallel planes at $x = 0$ and $x = a$, each mirror image of the charge in one plane is in turn reflected in the other, leading to the following series for the potential distribution:

$$\phi(x,y,z) = q \sum_{n=-\infty}^{\infty} \{([x-d+2na]^2 + y^2 + z^2)^{-\frac{1}{2}} - ([x+d+2na]^2 + y^2 + z^2)^{-\frac{1}{2}}\} \qquad [11\cdot116]$$

Where there is a charge distribution $\rho(x_o,y_o,z_o)$ between the two plates this series becomes

$$\phi(x,y,z) = \sum_{n=-\infty}^{\infty} \int \{([x-x_o+2na]^2 + [y-y_o]^2 + [z-z_o]^2)^{-\frac{1}{2}} - ([x+x_o+2na]^2 + [y-y_o]^2 + [z-z_o]^2)^{-\frac{1}{2}}\} \rho(x_o,y_o,z_o)\, dx_o\, dy_o\, dz_o \qquad [11\cdot117]$$

If, instead of being near a plane conductor, the charge q is at a distance h from the center of a conducting sphere of radius R, the charge induced in the surface is equivalent to an image charge $-qR/h$ at a distance

R^2/h from the center on a line joining the charge and the center of the sphere. The potential distribution can be determined from the charge and its image. The potentials due to the presence of two spheres, or of a continuous charge distribution, can be determined by applying the same type of reasoning employed for the plane configuration.

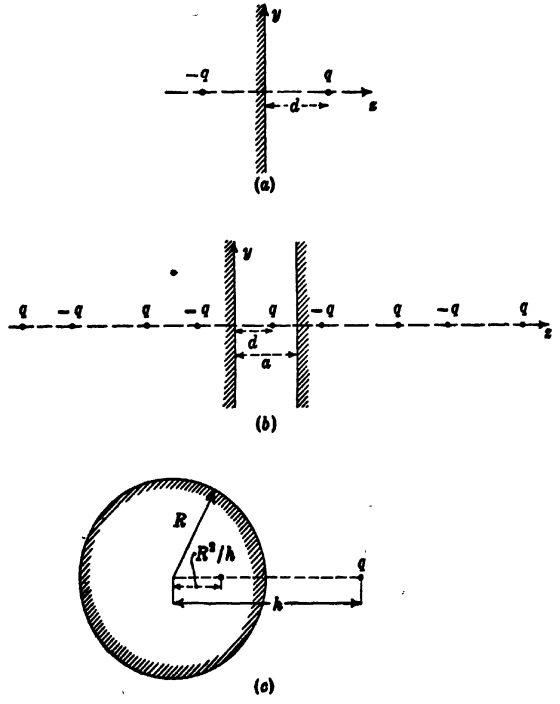


FIG. 11·25. Electric Images (*a*) at a Plane, (*b*) at Two Parallel Planes, and (*c*) at a Sphere.

When a conductor encloses charges it induces upon its inner surface a charge which is equal to the enclosed charge and opposite in sign. Thus if a line charge having a density of λ units of charge per unit length lies along the axis of a circular cylinder of radius R, the charge density induced on the inner surface has a value $\sigma = -\lambda/(2\pi R)$. If this charge is uniformly distributed over a beam of cross-section radius r_o, the potential on the axis of the beam will be

$$\phi(z,0) = \lambda\left(2\log\frac{R}{r_o} + 1\right) \qquad [11\cdot118]$$

whereas in the free space around the beam it will be

$$\phi(z,r) = 2\lambda \log \frac{R}{r} \qquad r_o < r < R \qquad [11\cdot119]$$

as can be readily calculated with the aid of Eq. 11·114.

These few examples are given to indicate the way in which the method of images can be used in estimating the effect of space charge in a given electron-optical configuration.

REFERENCES

1. J. JEANS, *Mathematical Theory of Electricity and Magnetism*, Cambridge University Press, 1933.
2. F. OLLENDORF, *Die Potentialfelder der Elektrotechnik*, J. Springer, Berlin, 1932.
3. E. P. ADAMS, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, Smithsonian Institution, Washington, 1922.
4. E. BRÜCHE and O. SCHERZER, *Geometrische Elektronenoptik*, J. Springer, Berlin, 1934.
5. E. JAHNKE and F. EMDE, *Funktionentafeln*, B. G. Teubner, Leipzig, 1933.
6. S. BERTRAM, "Calculation of axially symmetric fields," *J. Applied Phys.*, Vol. 13, pp. 496–502, August 1942.
7. S. BERTRAM, "Determination of the axial potential distribution in axially symmetric electrostatic fields," *Proc. Inst. Radio Engrs.*, Vol. 28, pp. 418–420, 1940.
8. M. KORMES, "Numerical solution of the boundary value problem for the potential equation by means of punched cards," *Rev. Sci. Instruments*, Vol. 14, pp. 248–250, August 1943.
9. G. H. SHORTLEY and R. WELLER, "The numerical solution of Laplace's equation," *J. Applied Phys.*, Vol. 9, pp. 334–348, 1938.
10. M. BOWMAN-MANIFOLD and F. H. NICOLL, "Electrolytic field-plotting trough for circularly symmetric systems," *Nature*, Vol. 140, p. 39, July 1938.
11. E. GUNDERT, "The aperture defect of electrostatic electron lenses," *Telefunkenröhre*, No. 19–20, pp. 61–98, March 1941.

CHAPTER 12

ELECTRON TRAJECTORY TRACING

12·1. The Ray Equation. The differential equation governing the path of an electron (or any charged particle) in an electrostatic field is most readily derived from the least-action principle:

$$\delta \int_a^b mv \cdot ds = 0 \qquad [12\cdot1]$$

The path of the electron between the terminal points a and b is such that the integral of its momentum over the path is a minimum[1] compared with the integral of the momentum over any neighboring path. The momentum of the electron is here at all points proportional to the square root of the electric potential ϕ, the origin of the potential scale having been chosen so that the sum of the kinetic and potential energies of the electron just vanishes.[2] Since the increment in any path length ds may be expressed in terms of the corresponding increment in the x-coordinate, dx, and the components of the slope of the path relative to the x-axis, $dy/dx = y'$ and $dz/dx = z'$:

$$ds = (1 + y'^2 + z'^2)^{\frac{1}{2}}\, dx$$

Equation 12·1 may also be written

$$\delta \int_a^b \phi^{\frac{1}{2}}(1 + y'^2 + x'^2)^{\frac{1}{2}}\, dx = \delta \int_a^b F(x,y,z;y',z')\, dx = 0 \qquad [12\cdot2]$$

The ray equations are the *Euler equations*[3] of this variation principle:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0 \qquad \frac{\partial F}{\partial z} - \frac{d}{dx}\frac{\partial F}{\partial z'} = 0 \qquad [12\cdot3]$$

The solutions of Eqs. 12·3,

$$y = y(x) \qquad z = z(x)$$

which satisfy the boundary conditions

$$y(x_a) = y_a \qquad z(x_a) = z_a \qquad y(x_b) = y_b \qquad z(x_b) = z_b$$

fulfil also the variation principle of Eqs. 12·1 and 12·2.

[1] Or, more generally, assumes a stationary value.

[2] Here, as well as in the succeeding chapters up to Chapter 18, the electron velocities are assumed to be small enough to render relativistic corrections negligible.

[3] See Woods, reference 1, pp. 317–331.

Assume now that the xy-plane is a plane of symmetry. Then the component normal to this plane of the force acting on a charged particle within it is zero. Hence a particle traveling to begin with in the plane of symmetry remains in it, its path being governed by the first equation in Eqs. 12·3.[4] This, after evaluation of the derivatives and transposing, becomes:

$$y'' = \frac{(1 + y'^2)}{2\phi}\left[\frac{\partial \phi}{\partial y} - y'\frac{\partial \phi}{\partial x}\right] \qquad [12\cdot4]$$

This equation may also be derived directly from Newton's second law of motion applied to a charged particle in an electric field. Thus, let the acceleration be resolved into two components, one in the direction of travel of the electron and the other normal to it. Then the normal component of the acceleration is v^2/R, R being the radius of curvature of the path. The Newtonian equation for this component becomes

$$\frac{mv^2}{R} = -eE_n \qquad [12\cdot5]$$

E_n is the component of the electric field in the direction of the normal to the path of the electron. The potential gradient, and hence the field component, in the direction of the normal to the path is given by

$$E_n = -\frac{\partial \phi}{\partial n} = -\left[-\frac{y'}{(1 + y'^2)^{1/2}}\frac{\partial \phi}{\partial x} + \frac{1}{(1 + y'^2)^{1/2}}\frac{\partial \phi}{\partial y}\right] \qquad [12\cdot6]$$

If, furthermore, familiar expressions for the curvature of the path are utilized:[5]

$$\frac{1}{R} = \frac{y''}{(1 + y'^2)^{3/2}} \qquad [12\cdot7]$$

and for the kinetic energy of the electron:

$$mv^2 = 2e\phi \qquad [12\cdot8]$$

Eq. 12·5 passes over into Eq. 12·4.

Equation 12·4 holds, of course, also for electron paths in the symmetry planes (meridional planes) of an axially symmetric system. Here it is convenient to express it in terms of the conventional coordinates r,z:

$$r'' = \frac{(1 + r'^2)}{2\phi}\left[\frac{\partial \phi}{\partial r} - r'\frac{\partial \phi}{\partial z}\right] \qquad [12\cdot9]$$

[4] For the equations for " skew " rays in axially symmetric systems, see Chapter 15, Eqs. 15·54 and 15·58.

[5] See Adams, reference 2, Formula 2.222.

If rays are considered which form very small angles with the axis and whose distance from the axis, also, remains small throughout the electrostatic field, ϕ may be replaced by the expansion in Eq. 11·75:

$$\phi = \sum_{n=0}^{\infty} (-1)^n \frac{\Phi^{(2n)}}{(n!)^2} \left(\frac{r}{2}\right)^{2n}$$

and all the terms in Eq. 12·9 of higher order than the first in r and r' can be dropped. Thus the *paraxial* ray equation

$$r'' = -\frac{\Phi'}{2\Phi} r' - \frac{\Phi''}{4\Phi} r \qquad [12\cdot10]$$

is obtained, Φ signifying the potential along the axis ($r = 0$). A very useful modification of Eq. 12·10, involving only the potential along the axis and its first derivative, is obtained by introducing the new variable $R = r \cdot \Phi^{1/4}$:[6]

$$R'' = -\frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 R \qquad [12\cdot11]$$

In analogy to Eq. 12·10 the motion of electrons in a " two-dimensional " field near the symmetry plane $y = 0$ is given by

$$y'' = -\frac{\Phi'}{2\Phi} y' - \frac{\Phi''}{2\Phi} y \qquad [12\cdot12]$$

It is always possible to reduce a linear differential equation of the second order such as Eq. 12·10 to a first-order equation. Making the substitution

$$c = -\frac{r'}{r} \qquad [12\cdot13]$$

results in

$$c' = c^2 - \frac{\Phi'}{2\Phi} c + \frac{\Phi''}{4\Phi} \qquad [12\cdot14]$$

The quantity c is known as the *convergence* of the ray. As shown in Fig. 12·1, it is the reciprocal of the distance between the reference plane normal to the axis, $z = z_P$, and the intersection of the tangent to the path with the axis. Once Eq. 12·14 has been integrated from the initial point (z_o, r_o) onward, the actual path $r(z)$ is obtained by a further quadrature:

$$r = r_o \cdot \exp\left[-\int_{z_o}^{z} c \cdot dz\right] \qquad [12\cdot15]$$

[6] See Scherzer, reference 3.

The convergence variable is particularly convenient in studying the focusing of electron pencils leaving a cathode, or surface which emits electrons with very small initial velocities. If these initial velocities were exactly zero, all electrons leaving a given point of the cathode would follow the same path. In particular, if this point is the intersection of

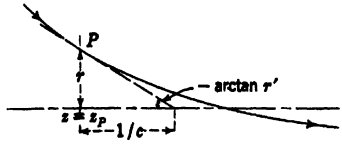


FIG. 12·1. Definition of the "Convergence" c.

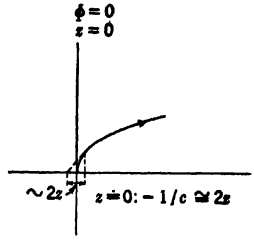


FIG. 12·2. Initial Divergence of Electrons Leaving a Cathode.

the axis of an axially symmetric system with the cathode, the electron paths will coincide with the axis. Assume, on the other hand, that the electrons have infinitesimal lateral initial velocity components. They will then, in the immediate neighborhood of the cathode, describe parabolas with their vertices at the point of origin on the cathode, since here the electric field can be regarded as uniform. Since for a parabola with a vertex at the origin (Fig. 12·2) the slope at any point is equal to half the ratio of the coordinates r/z, the convergence of the path in this region is given, to a good approximation, by

$$c = -\frac{1}{2z} \qquad [12\cdot16]$$

Thus, although the convergence itself becomes infinite at the cathode ($z = 0$), the variable

$$b = c + \frac{1}{2z} \qquad [12\cdot17]$$

remains finite. With b as dependent variable, the ray equation takes on the form

$$b' = b^2 - b\left(\frac{1}{z} + \frac{\Phi'}{2\Phi}\right) + \frac{\Phi''}{4\Phi} + \frac{1}{2z}\left(\frac{\Phi'}{2\Phi} - \frac{1}{2z}\right) \qquad [12\cdot18]$$

and the separation of the ray from the axis becomes

$$r = r_0\left(\frac{z}{z_0}\right)^{1/4} \exp\left[-\int_{z_0}^{z} b\,dz\right] \qquad [12\cdot19]$$

Here r_o is the separation from the axis prescribed for a particular axial coordinate z_o.

In order to integrate Eq. 12·18 it is necessary to know the initial values of b and b'. For this purpose expand b and Φ in powers of z:[7]

$$b = b_o + b_o' z + \cdots \qquad [12\cdot20]$$

$$\Phi = \Phi_o' z + \frac{\Phi_o'' z^2}{2} + \frac{\Phi_o''' z^3}{6} + \cdots \qquad [12\cdot21]$$

and substitute the series in Eq. 12·18. In order that Eq. 12·18 should be satisfied for all values of z it is then necessary that the coefficients of the several powers of $z(z^{-1}, z^0, z^1, \cdots)$ fulfil it individually. The first two of the resulting group of equations lead to the initial values

$$b_o = \frac{\Phi_o''}{4\Phi_o'} \qquad [12\cdot22]$$

$$b_o' = -\frac{1}{40}\left(\frac{\Phi_o''}{\Phi_o'}\right)^2 + \frac{2}{15}\frac{\Phi_o'''}{\Phi_o'} \qquad [12\cdot23]$$

If, as is frequently the case, only the point of convergence on the axis, and not the actual ray path, is sought, it is not necessary to carry out the quadrature indicated in Eq. 12·19. By starting with the initial values given by Eqs. 12·22 and 12·23, Eq. 12·18 is integrated to some point beyond the refracting fields, where the potential is substantially constant. Let this point have the axial coordinate z_b and let $b(z_b) = b_b$. Then the point at which the ray intersects the axis will be given by

$$z_i = z_b - \frac{r_b}{r_b'} = z_b + \frac{2z_b}{2b_b z_b - 1} = z_b \frac{2b_b z_b + 1}{2b_b z_b - 1} \qquad [12\cdot24]$$

The expansion Eq. 12·21 is useful also in determining the initial value of r'' for electrons leaving off-axis points of a cathode with zero initial velocity. Here Eq. 12·10 is most convenient. Since the initial slope of the ray is normal to the surface of the cathode, $r_o' = 0$ for a flat cathode. In this special case, furthermore, all even derivatives of the axial potential vanish at $z = 0$, since $\phi(0,r) = 0$. Substitution of Eq. 12·21 in Eq. 12·10 thus leads simply to

$$r_o'' = -\frac{1}{6}\frac{\Phi_o'''}{\Phi_o'} r_o \qquad [12\cdot25]$$

12·2. Analytical Solution of the Ray Equation. Only in a very few special cases is it possible to write down a closed expression for the electron paths determined by Eqs. 12·4, 12·10, or 12·12. The simplest of

[7] The subscript o here refers to values at $z = 0$, not at $z = z_o$.

all cases is, of course, that of the uniform electric field (Fig. 12·3):

$$\frac{\partial \phi}{\partial x} = -E \qquad \frac{\partial \phi}{\partial y} = 0 \qquad \phi = -E(x - x_o) \qquad [12\cdot26]$$

In this case Eq. 12·4 becomes:

$$y'' = -\frac{(1 + y'^2)}{2(x - x_o)} y' \qquad [12\cdot27]$$

This equation has the general solution:

$$y = B + 2A^{\frac{1}{2}}(x - x_o - A)^{\frac{1}{2}} \qquad [12\cdot28]$$

This describes a family of parabolas whose vertex lies at $x = x_o + A$,

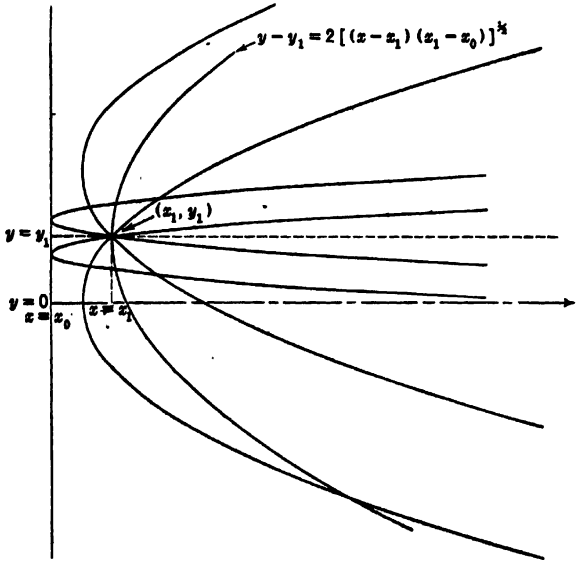


FIG. 12·3. Parabolic Electron Paths in Uniform Electric Field.

$y = B$. If the path passes through the point x_1, y_1 and its slope at that point is y_1' (Fig. 12·3), the constants A and B are given by

$$A = (x_1 - x_o)\frac{y_1'^2}{1 + y_1'^2} \qquad B = y_1 - \frac{2y_1'}{1 + y_1'^2}(x_1 - x_o) \qquad [12\cdot29]$$

If, in particular, the point x_1, y_1 corresponds to the vertex of the parabola,

$$y - y_1 = 2(x - x_1)^{\frac{1}{2}}(x_1 - x_o)^{\frac{1}{2}} \qquad [12\cdot30]$$

The same result may be obtained directly from the equations of motion of the electron in a uniform field:

$$y - y_1 = v_y t = \left(\frac{2eV_y}{m}\right)^{1/2} t$$

$$x - x_1 = -\frac{eE}{2m} t^2 \qquad [12\cdot31a]$$

Here v_y is the velocity component in the y-direction and eV_y the kinetic energy corresponding to this velocity component; t is the time. By eliminating t from the two equations,

$$y - y_1 = v_y \left(\frac{2m(x - x_1)}{-eE}\right)^{1/2} = 2\left(\frac{V_y}{-E}\right)^{1/2} (x - x_1)^{1/2} \qquad [12\cdot31b]$$

Since at the point (x_1,y_1), $\phi = V_y$, as the kinetic energy connected with motion in the x-direction vanishes at the vertex of the parabola, $x_1 - x_o = -V_y/E$ and

$$y - y' = 2(x_1 - x_o)^{1/2}(x - x_1)^{1/2}$$

as before.

Another case which can be treated analytically, and which is of some interest because of its analogy to the homogeneous magnetic field, is the electrostatic field which increases or decreases exponentially along the axis of an axially symmetric configuration:

$$\Phi(z) = C \cdot e^{bz} \qquad [12\cdot32a]$$

$$\phi(z,r) = C \cdot e^{bz} \cdot J_0(br) \qquad [12\cdot32b]$$

Here $J_0(x)$ is the Bessel function of zero order. Equation 12·32*b* follows from Eq. 12·32*a* if the latter is substituted in Eq. 11·75.

Substituting Eq. 12·32 in Eq. 12·11 immediately leads to the solution

$$R = A_1 \sin \frac{\sqrt{3}}{4} b(z - z_o) \qquad [12\cdot33a]$$

$$r = Ae^{-bz/4} \sin \frac{\sqrt{3}}{4} b(z - z_o) \qquad [12\cdot33b]$$

The path of the electron is thus a sine function of exponentially increas-

ing or exponentially decreasing amplitude, depending on whether the field is retarding or accelerating it, respectively (Fig. 12·4).

A few more problems capable of solution by analytical methods might be cited. Nevertheless, without any question, this approach is extremely limited.

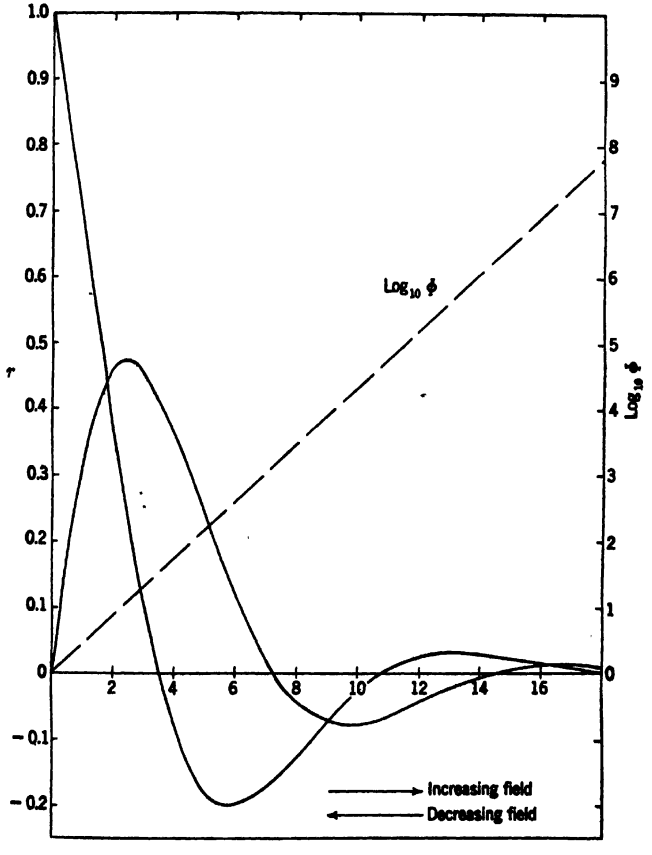


FIG. 12·4. Paths in Exponentially Increasing or Decreasing Field.

12·3. Numerical Methods; Differences. The method of numerical integration furnishes a perfectly general way of determining electron paths in given electric fields. As numerical methods are of great practical importance in many different aspects of electron optics, it is ad-

visable to begin with the simplest case, that is, the numerical integration of a known function $y = f(x)$. The procedure consists in listing values of y for equally spaced values of x and then forming the differences

$$\begin{aligned} \Delta_1 y_i &= y_i - y_{i-1} \\ \Delta_2 y_i &= \Delta_1 y_i - \Delta_1 y_{i-1} \\ \Delta_3 y_i &= \Delta_2 y_i - \Delta_2 y_{i-1} \\ &\cdots\cdots\cdots\cdots \end{aligned} \qquad [12\cdot34]$$

This procedure yields the following array:

$x_o - 4h$	y_{-4}				
		$\Delta_1 y_{-3}$			
$x_o - 3h$	y_{-3}		$\Delta_2 y_{-2}$		
		$\Delta_1 y_{-2}$		$\Delta_3 y_{-1}$	
$x_o - 2h$	y_{-2}		$\Delta_2 y_{-1}$		$\Delta_4 y_o$
		$\Delta_1 y_{-1}$		$\Delta_3 y_o$	
$x_o - h$	y_{-1}		$\Delta_2 y_o$		$\Delta_4 y_1$
		$\Delta_1 y_o$		$\Delta_3 y_1$	
x_o	y_o		$\Delta_2 y_1$		$\Delta_4 y_2$
		$\Delta_1 y_1$		$\Delta_3 y_2$	
$x_o + h$	y_1		$\Delta_2 y_2$		$\Delta_4 y_3$
		$\Delta_1 y_2$		$\Delta_3 y_3$	
$x_o + 2h$	y_2		$\Delta_2 y_3$		.
		$\Delta_1 y_3$		.	
$x_o + 3h$	y_3		.		
		.			
.	.				

The intervals should be chosen small enough that the highest-order differences are so small that they have a negligible effect on the final result. If the function $y = f(x)$ has been determined experimentally, rather large random variations in the higher differences must be expected, owing to errors of measurement. These may often be compensated by adjusting the measured function values until the higher differences run smoothly.

Once the differences have been tabulated, the integral over any interval h is expressible in terms of the function and its differences as follows:

$$\int_{x_o-h}^{x_o} y\,dx = h\left(y_o - \frac{1}{2}\Delta_1 y_o - \frac{1}{12}\Delta_2 y_o - \frac{1}{24}\Delta_3 y_o - \frac{19}{720}\Delta_4 y_o - \frac{3}{160}\Delta_5 y_o - \cdots\right) \qquad [12\cdot35]$$

Normally it is wise, from the point of view of economy of effort, to choose intervals small enough that the series may be terminated with the third difference. Then the procedure is equivalent to fitting a third-order curve through the point (x_o,y_o) and the three preceding points and evaluating the area underneath the curve between the ordinates $x = x_o - h$ and $x = x_o$. If the third-order curve is represented by

$$y = y_o + a(x - x_o) + b(x - x_o)^2 + c(x - x_o)^3 \qquad [12\cdot36]$$

the several values of y and the differences become

$$\begin{array}{lll} y_o = y_o & \Delta_1 y_o = ah - bh^2 + ch^3 & \Delta_2 y_o = 2bh^2 - 6ch^3 \\ y_{-1} = y_o - ah + bh^2 - ch^3 & \Delta_1 y_{-1} = ah - 3bh^2 + 7ch^3 & \Delta_2 y_{-1} = 2bh^2 - 12ch^3 \\ y_{-2} = y_o - 2ah + 4bh^2 - 8ch^3 & \Delta_1 y_{-2} = ah - 5bh^2 + 19ch^3 & \\ y_{-3} = y_o - 3ah + 9bh^2 - 27ch^3 & & \Delta_3 y_o = 6ch^3 \end{array}$$

The first four equations may be solved for the constant coefficients in Eq. 12·36, yielding:

$$ch^3 = \frac{\Delta_3 y_o}{6} \qquad bh^2 = \frac{\Delta_2 y_o}{2} + \frac{\Delta_3 y_o}{2} \qquad ah = \Delta_1 y_o + \frac{\Delta_2 y_o}{2} + \frac{\Delta_3 y_o}{3}$$

Integrating Eq. 12·36 between $x_o - h$ and x_o leads to the first four terms on the right of Eq. 12·35:

$$\int_{x_o-h}^{x_o} y\,dx = y_o h - \frac{ah^2}{2} + \frac{bh^3}{3} - \frac{ch^4}{4}$$

$$= h\left(y_o - \frac{\Delta_1 y_o}{2} - \frac{\Delta_2 y_o}{12} - \frac{\Delta_3 y_o}{24}\right)$$

The third-order curve may, of course, equally well be fitted to any other four successive points adjoining the interval considered. Thus, if the points beyond and including (x_o,y_o) are used, the series

$$\int_{x_o}^{x_o+h} y\,dx = h\left(y_o + \frac{1}{2}\Delta_1 y_1 - \frac{1}{12}\Delta_2 y_2 + \frac{1}{24}\Delta_3 y_3 - \cdots\right) \qquad [12\cdot37]$$

results. Although certain intermediate selections of "fitting points" lead to even more rapidly converging series, Eqs. 12·35 and 12·37 are particularly convenient and adequate for all practical purposes. Examples for the application of the method may be taken from any book of mathematical tables.

Assume next that the differential equation

$$\frac{dy}{dx} = f(x,y) \qquad [12\cdot38]$$

is to be integrated, the initial value of y, at $x = x_o$, being y_o. Then

$$y = y_o + \int_{x_o}^{x} f(x,y)\,dx \qquad [12{\cdot}39]$$

The procedure is again to tabulate x, $y' = f(x,y)$, and y for equal intervals in x as follows:

x_o	$y'_o = f(x_o,y_o)$				y_o
$x_o + h$	$y'_1 = f(x_o + h, y_1)$	$\Delta_1 y'_1$			y_1
$x_o + 2h$	$y'_2 = f(x_o + 2h, y_2)$	$\Delta_1 y'_2$	$\Delta_2 y'_2$		y_2
$x_o + 3h$	$y'_3 = f(x_o + 3h, y_3)$	$\Delta_1 y'_3$	$\Delta_2 y'_3$	$\Delta_3 y'_3$	y_3
.....	.....	.....	.....	.....	.....
$x_o + (n-1)h$	$y'_{n-1} = f(x_o + [n-1]h, y_{n-1})$	$\Delta_1 y'_{n-1}$	$\Delta_2 y'_{n-1}$	$\Delta_3 y'_{n-1}$	y_{n-1}
$x_o + nh$	$y'_n = f(x_o + nh, y_n)$	$\Delta_1 y'_n$	$\Delta_2 y'_n$	$\Delta_3 y'_n$	y_n

Suppose, to begin with, that y_o, y_1, $\cdots y_{n-1}$ are known and that the integration is to be carried forward from this point. To begin with the difference $y_n - y_{n-1}$ is guessed; for example, by comparison with the preceding differences $y_{n-1} - y_{n-2}$, $y_{n-2} - y_{n-3}$, etc. This is substituted in Eq. 12·38 to obtain a value of y'_n. After differences have been formed, the quantity $y_n - y_{n-1}$ is evaluated anew from

$$y_n - y_{n-1} = h\left(y'_n - \frac{1}{2}\Delta_1 y'_n - \frac{1}{12}\Delta_2 y'_n - \frac{1}{24}\Delta_3 y'_n - \cdots\right) \qquad [12{\cdot}40]$$

If the value thus obtained agrees with that originally guessed, the value y_n is established and the integration may proceed to the next interval. However, if this is not so, the newly found value of y_n (or some intermediate value) is substituted in Eq. 12·38, and the resulting value of y'_n and the corresponding differences are used to redetermine the value of $y_n - y_{n-1}$ by Eq. 12·40. This is continued until agreement between the assumed value of $y_n - y_{n-1}$ and that determined by Eq. 12·40 is obtained. In practice this rarely requires more than two trials.

In initiating the integration the higher differences of the integrand are not available. Hence, for the first interval, y_1 may be approximated by

$$y_1 - y_o = h\left(y'_1 - \frac{\Delta_1 y'_1}{2}\right)$$

For the next interval the second-order difference will be available so that the first two differences in Eq. 12·40 can be utilized. When, finally, y_3, together with the corresponding differences, has been established, y_1 may be redetermined, making use of a series of the type given in Eq. 12·37. The resulting change will require a redetermination also

of y_2 and y_3, as well as of y_2' and y_3'. If the several differences have been altered considerably as a result, a third determination of y_1, y_2, and y_3 will be needed. In short, the process must be continued until these

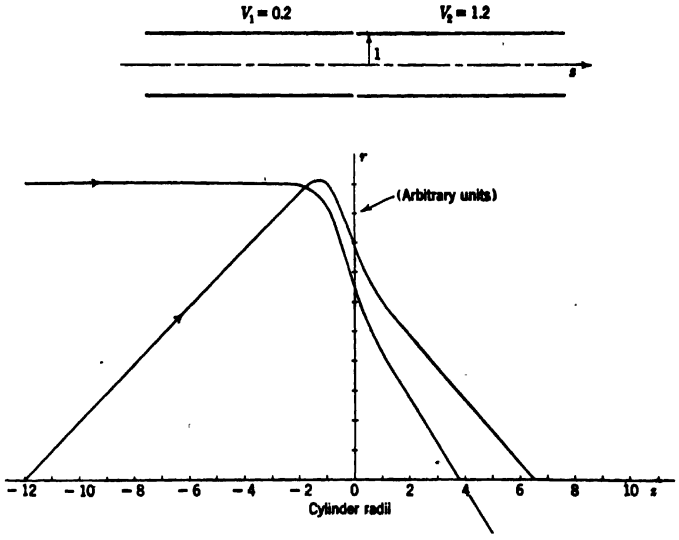


FIG. 12·5. Paraxial Rays in the Field between Two Equidiameter Coaxial Cylinders.

values are consistent with those determined by Eqs. 12·37 and 12·40. Small useful variations in this procedure will occur to any one who has occasion to carry out numerical integrations of differential equations.

In a second-order equation, such as Eqs. 12·4, 12·10, or 12·11,

$$y'' = f(x,y,y') \qquad [12\cdot41]$$

the integration is broken up into two steps:

$$y' = y_o' + \int_{x_o}^{x} f(x,y,y')\,dx \qquad y = y_o + \int_{x_o}^{x} y'\,dx \qquad [12\cdot42]$$

the initial values y_o and y_o' at $x = x_o$ being given. Values of y and y' are guessed and substituted in $f(x,y,y')$. This, by Eq. 12·42, is used to find a value for y' and the latter value of y' used to determine y, applying difference series of the character of those given in Eqs. 12·37 and 12·40. Again, the procedure is continued until the guessed and the integrated values of y' and y agree. Electron paths in the electrostatic multiplier structure shown in Fig. 11·4 and indicated there by heavy lines were determined by numerical integration of Eq. 12·4.

It is to be noted that the *paraxial* equations (12·10, 12·11, and 12·12)

are homogeneous in the dependent variables r, R, and y, respectively. Accordingly all paths with the same initial values of r'_o/r_o, R'_o/R_o, or y'_o/y_o are quantitatively similar. Thus, although, in fact, these equations are valid only for rays very close to the axis, it is legitimate, in their integration, to give the initial separation from the axis or the initial slope any arbitrary convenient value, such as unity. The correct path for a very small initial separation from the axis or initial slope is obtained by multiplying all the separations of the path calculated by an appropriate constant factor. Figure 12·5 shows paraxial rays traced through the electrostatic field between two cylinders, calculated by the integration of Eq. 12·10. It is well to note that in all cases where the potential distribution is determined experimentally by electrolytic plotting, Eq. 12·11 is greatly to be preferred to Eq. 12·10. Equation 12·11 involves only the first derivative of the axial potential, Φ', which can be obtained with much less effort and much greater accuracy than the second derivative, Φ''.

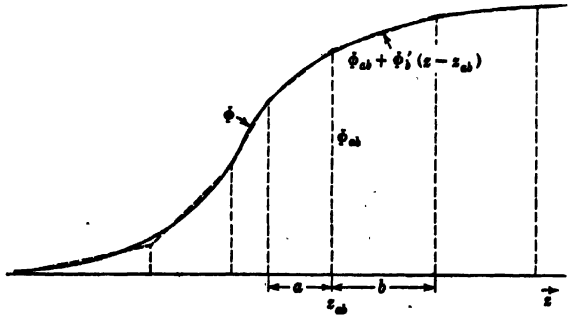


FIG. 12·6. Approximation of the Axial Potential Curve by a Broken Line.

12·4. Path Integration in Approximated Fields (Gans).[8] An approximation to the path of an electron in an arbitrary field can be obtained if the actual field is replaced by one which closely resembles it, yet permits analytical integration of the path. This procedure is often well adapted for the calculation of paraxial rays in axially symmetric systems. In its simplest and most useful form it involves replacing the variation of the axial potential by a broken line, approximating as closely as possible in a prescribed number of steps the actual potential curve (Fig. 12·6). Thus the true field is replaced by a sequence of regions of uniform field, the degree of approximation improving with the number of steps.

[8] See reference 4.

The integration of Eq. 12·10 in the " step field " may be divided into two parts: (1) the integration from one end of a field segment to the other and (2) the integration over an infinitesimal distance enclosing the " break point."

Within any segment $\Phi'' = 0$, and Eq. 12·10 may be written

$$\Phi^{1/4} r'' + \frac{\Phi'}{2\Phi^{1/2}} r' = 0 \qquad [12\cdot43]$$

which is integrated by

$$\Phi^{1/4} r' = C \qquad [12\cdot44]$$

For example, for section b, $C = C_b$ is a constant determined by the values of r' and Φ at the initial point of the segment, r'_{ab+} and Φ_{ab}. The subscript ab refers to the junction between segments a and b and the plus sign to the fact that the value of r' to the right of the junction between a and b is to be taken.

Since

$$\Phi_b = \Phi_{ab} + (z_b - z_{ab})\Phi'_b \qquad [12\cdot45]$$

Equation 12·44 may be further integrated to yield

$$r_b = r_{ab} + \int_{z_{ab}}^{z_b} \frac{C_b\, dz}{(\Phi_{ab} + [z - z_{ab}]\Phi'_b)^{1/2}} \qquad [12\cdot46]$$

$$= r_{ab} + \frac{2C_b}{\Phi'_b}(\Phi_b^{1/2} - \Phi_{ab}^{1/2}) \quad \text{for} \quad \Phi'_b \neq 0$$

For $\Phi'_b = 0$ (a field-free region) the solution simply becomes, obviously,

$$r_b = r_{ab} + r'_{ab+}(z_b - z_{ab}) \qquad [12\cdot47]$$

At the junction (for example, at $z = z_{ab}$) Φ'' becomes infinite, whereas Φ, Φ', r, and r' all remain finite. Since, furthermore, Φ and r are continuous as well, they may be treated as constants. Hence the integration of Eq. 12·10 across the junction

$$\int_{z_{ab-}}^{z_{ab+}} r''\, dz = -\int_{z_{ab-}}^{z_{ab+}} \left(\frac{\Phi'}{2\Phi} r' + \frac{\Phi''}{4\Phi} r\right) dz \qquad [12\cdot48]$$

reduces to

$$r'_{ab+} - r'_{ab-} = \frac{\Phi'_{ab-} - \Phi'_{ab+}}{4\Phi_{ab}} r_{ab} \qquad [12\cdot49]$$

With the aid of Eqs. 12·46, 12·47, and 12·49 it is possible to integrate any paraxial electron path through a step field provided that the initial

height of incidence and slope of the path are given. An application of the method to the determination of the focal properties of aperture lenses is given in sections 13·7 and 13·8.

It is also possible to integrate analytically the paths of electrons in fields which have been approximated by a sequence of parabolic arcs, as has been shown by Gans[9] and Recknagel.[10] Although this makes it possible to reduce, for a given accuracy, the number of steps, the curve fitting and the formulas for the paths within the field segments become considerably less simple.

12·5. Graphical Methods. Although the numerical methods described in the two preceding sections are most important in the determination of the focusing properties of electron lenses, discussed further in Chapters 13 and 15, graphical methods have great utility in the tracing of paths in electrostatic fields, whose primary purpose is not image formation. Fields of this character occur in electrostatic multipliers and various beam-forming systems.

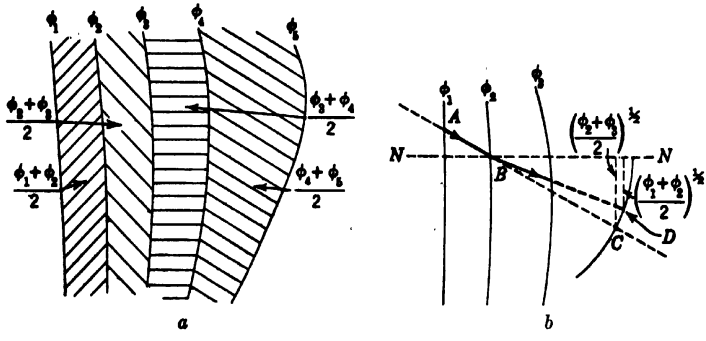


FIG. 12·7. Graphic Path Plotting by Snell's Law.

The most obvious graphical method is based on a direct application of Snell's law of refraction to the equipotential surfaces. Every one of a series of such surfaces, spaced in equal intervals of potential, is considered as separating two homogeneous regions at potentials equal to the average of those of the bounding surfaces (Fig. 12·7*a*). The incident ray AB (Fig. 12·7*b*) on such a surface (ϕ_2) is continued to a point C from which the perpendicular dropped on the normal NN is equal (or proportional) in length to $([\phi_2 + \phi_3]/2)^{1/2}$. Then the point D on the circle through C about B, for which the perpendicular dropped to NN

[9] See reference 4.
[10] See reference 5.

is equal to $([\phi_1 + \phi_2]/2)^{\frac{1}{2}}$, determines the refracted ray BD. This process is repeated for each surface, approximating the path between surfaces by straight line segments.

A graphical method more suitable, in particular, where the electron path makes but small angles with the equipotentials is the *circle method* based on Eqs. 12·5 and 12·8. From these it follows that

$$R = -\frac{2\phi}{E_n} = \frac{\phi_2 + \phi_1}{\phi_2 - \phi_1} D \qquad [12\cdot50]$$

Here R is the radius of curvature of the ray between equipotentials $\phi = \phi_1$ and $\phi = \phi_2$. D is the distance between the point of incidence on equipotential 1 and the intersection between the normal to the path at 1 with the tangent to equipotential 2, as shown in Fig. 12·8. The path is thus approximated by a series of circular segments joined smoothly together.

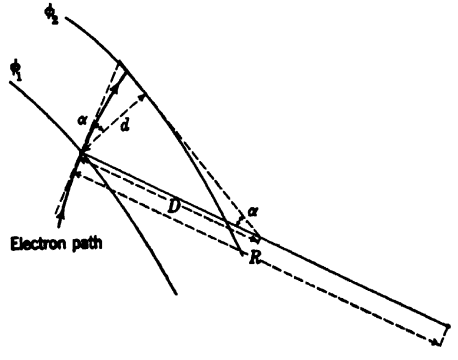


FIG. 12·8. "Circle Method" of Path Plotting.

In this connection it is of interest to derive an expression for the curvature of the path of an electron in a "two-dimensional" system as it leaves a surface at zero potential; for example, is emitted by a cathode.[11] If the normal to the surface at the point of emission is taken as the x-axis, the curvature is given simply by the second derivative of y with respect to x (Eq. 12·4):

$$y'' = \frac{1 + y'^2}{2\phi}\left[\frac{\partial\phi}{\partial y} - y'\frac{\partial\phi}{\partial x}\right]$$

At $x = 0$ this is indefinite, as $\phi = y' = \partial\phi/\partial y = 0$. However, differ-

[11] See Rajchman, reference 6.

entiating, in accord with Hospital's rule,[12] both numerator and denominator, and solving once more for y'' leads to

$$y'' = \frac{1}{3} \frac{\dfrac{\partial^2 \phi}{\partial y \partial x}}{\dfrac{\partial \phi}{\partial x}} \qquad [12\cdot51]$$

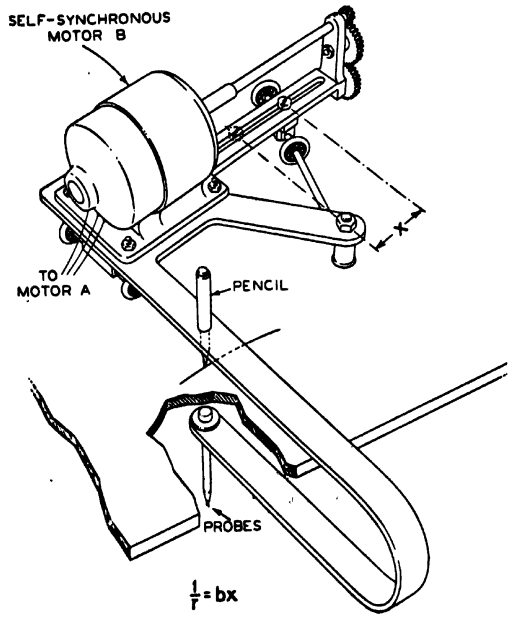


FIG. 12·9. Recording Carriage of Automatic Path Plotter. (Courtesy *Nature*, reference 8.)

Now, if $\psi(x,y) = \text{const}$ is the equation of the force line through the point of emission, according to Eq. 11·8:

$$\frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x} \qquad \frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y}$$

so that

$$y'' = \frac{1}{3} \frac{\dfrac{\partial^2 \psi}{\partial y^2}}{\dfrac{\partial \psi}{\partial y}} \qquad [12\cdot52]$$

[12] See Adams, reference 2, Formula 7.102.

The expression to the right of Eq. 12·52 is simply one-third of the curvature of the force line $\psi(x,y) = \text{const}$. The curvature of the ray at the point of emission is thus one-third as large as that of the force line through the same point.

12·6. Mechanical Path-Plotting Mechanisms. Equations 12·5 and 12·8, on which the circle method of path plotting is based, leading to

$$R = -\frac{2\phi}{E_n}$$

form also the starting point of path-plotting mechanisms, used directly in conjunction with the electrolytic tank, which have been described by Gabor[13] and D. B. Langmuir.[14]

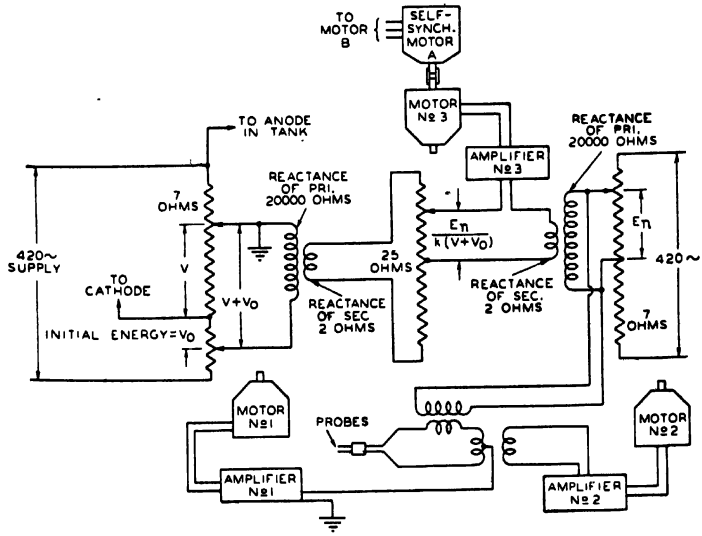


FIG. 12·10. Circuit of Automatic Path Plotter. (Courtesy *Nature*, reference 8.)

Figure 12·9 is a sketch of the moving carriage bearing both the two slightly separated probes and the recording pencil in Langmuir's device, which alone is completely automatic. The probes and the pencil are rigidly connected by a U-shaped bar. The path-plotting paper is placed on a plate of glass directly above the tank. The plane containing the two probes is maintained normal to the direction of motion of the carriage, and, as shown in the circuit diagram (Fig. 12·10), their differ-

[13] See reference 7.

[14] See reference 8.

ence of potential is used to measure the normal component of the field E_n. This is set automatically by motor 2; the mean potential of the two electrodes is indicated on another potentiometer by motor 1. Finally, motor 3 rotates until the setting on the central potentiometer is proportional to E_n/ϕ (in the figure $\phi = V + V_o$). The setting on the central potentiometer is translated into a corresponding setting of the axle of the rear pair of wheels by self-synchronous motors, so that the radius of curvature is always given by

$$R = -\frac{2\phi}{E_n}$$

The carriage as a whole is propelled by a small motor at a rate of approximately 1 millimeter per second.

12·7. The Rubber Model. For " two-dimensional " systems electron paths may be obtained directly, without a separate determination of the potential distribution between the electrodes, by much simpler means, the rubber model. The operation of this device is based on two circumstances:

1. A thin elastic membrane stretched over a model of an electrode configuration whose height is everywhere proportional to the potential energy of an electron in the immediate neighborhood adjusts itself in such fashion that its height at any point is proportional to the space potential at that point.
2. The projection on a horizontal plane of the path of a small sphere rolling without friction on this membrane under the influence of gravity is identical with the path of an electron traveling in a plane of symmetry of the two-dimensional electrode structure.

Both statements are valid only under restrictions which cannot be exactly fulfilled in practice. However, the conditions of their validity can readily be approximated sufficiently closely to make the rubber model an exceedingly effective means for determining electron paths in " two-dimensional " systems.

The first property of the rubber model follows from the fact that on an ideal membrane — that is, a two-dimensional body maintaining uniform tension over its entire area — with equal pressure (atmospheric pressure) on both sides every point is a saddle point. In other words, the sum of the reciprocals of the radii of curvature in two mutually perpendicular planes through the normal at the point considered vanishes (Fig. 12·11). If the surface at the point in question is horizontal, planes parallel to the xz- and yz-planes may be chosen as the appropriate

pair of perpendicular planes, and the sum of the reciprocal radii becomes

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0 \qquad [12\cdot53]$$

At points where the surface does not have a horizontal tangent plane this equation applies with an accuracy decreasing with the deviation of the tangent plane from the horizontal. In practice angles as large as 15 degrees do not lead to serious error.

Since z is proportional to the potential of the electrodes where the membrane is in contact with them and fulfils Laplace's equation in the region between them, the height of the membrane represents in fact the electric potential throughout.

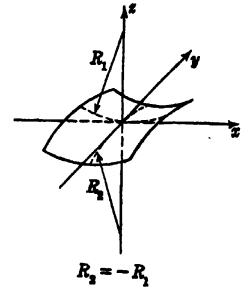


FIG. 12·11. Curvature of Membrane between Regions at Equal Pressure.

Consider now a small sphere, of mass m and radius a, rolling on this model. If the curvature of the membrane is everywhere negligible compared to the curvature of the sphere, the angular velocity about the center of the sphere ω and the linear velocity of translation v are related by

$$v = a\omega \qquad [12\cdot54]$$

If there is no friction except the static friction necessary to prevent the sphere from sliding, the kinetic energy of the sphere is given by

$$T = \frac{mv^2}{2} + \frac{I\omega^2}{2} = \left(m + \frac{I}{a^2}\right)\frac{v^2}{2} = \frac{7}{10}mv^2 \qquad [12\cdot55]$$

$I = \frac{2}{5}ma^2$ is the moment of inertia of the sphere about a diameter. The potential energy is

$$\text{P.E.} = mgz + \text{const} \qquad [12\cdot56]$$

where z is the height of the center of gravity measured from an arbitrary level and g is the acceleration of gravity. If the kinetic energy just vanishes at the level $z = z_o$,

$$\frac{7}{10}mv^2 = mg(z_o - z)$$

and

$$v = \text{const}\ (z_o - z)^{1/2} \qquad [12\cdot57]$$

Since the total momentum of the sphere is given by $\partial T/\partial v = \frac{7}{5}mv$,

the path of the center of gravity is governed, according to the principle of least action, by

$$\delta \int (z_o - z)^{1/4}\, ds = \delta \int (z_o - z)^{1/4} \left(\left[\frac{dz}{dx}\right]^2 + \left[\frac{dy}{dx}\right]^2 + 1 \right)^{1/4} dx = 0$$

If the slope of the membrane is small at all points, $(dz/dx)^2 \ll 1$, it is possible to write

$$\delta \int z_1^{1/4} (y'^2 + 1)^{1/4}\, dx = 0 \qquad z_1 = z_o - z$$

which is exactly the variation principle giving the path of an electron in the potential field $\phi \propto z_1$. Although this deduction, again, is not rigor-



FIG. 12-12. The Rubber Model for the Study of Electron Paths in Two-Dimensional Systems.

ous, the result is approximated sufficiently closely in practice to make this method of path plotting in two-dimensional systems highly useful.

Figure 12-12 shows a rubber model set up to indicate electron paths in a particular system. The membrane of surgical rubber stretched on a wooden frame has been placed over the electrode model made of strips of sheet aluminum cut to the right height and bent into proper shape. Horizontal arms with clamps are provided to press the rubber against

the electrodes corresponding to more positive potentials by means of similarly shaped strips of aluminum, the *top electrodes*. A light source

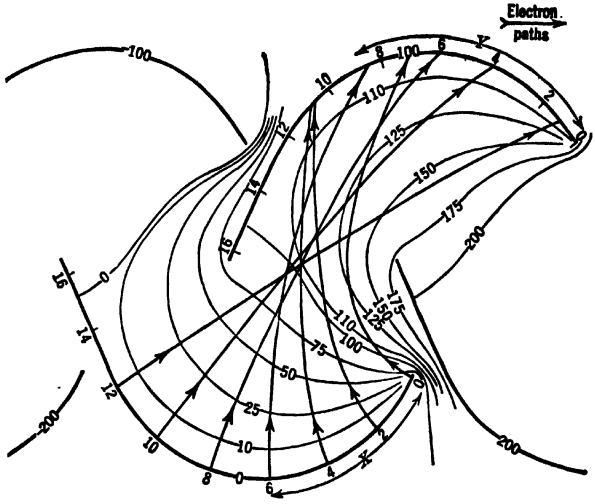


FIG. 12·13. Paths between Target Electrodes of Electrostatic Multiplier. (Courtesy *Proc. Inst. Radio Engrs.*, reference 10.)

shining through the plate glass base plate facilitates the checking of the model electrode positions underneath the rubber. Steel ball bearings $\frac{3}{16}$ inch in diameter conveniently represent the moving electrons. With the aid of an electromagnet they may be released at any point without an undesirable initial impact.

In order to record the electron paths in the model, a time-exposure photograph may be taken with a camera mounted above the model. The brightly illuminated moving spheres leave a clearly marked track on the sensitive plate. If, furthermore, the light source is intermittent — for example, a mercury arc fed by alternating current — this track will consist of short dashes, the length of each dash being proportional to the speed of the electron.[15]

The method of the rubber model has played a pre-eminent part in the development of the electrostatic electron multiplier. Figure 12·13[16] shows paths between two of the target electrodes of a particular configuration, these having been determined by the graphical method of path plotting described on p. 415 (Fig. 12·8). Figure 12·14[16] compares the positions of the terminal points of the electron paths on the target

[15] See Kleynen, reference 9.

[16] See Zworykin and Rajchman, reference 10.

electrode y as function of the position of their initial points on target electrode x determined by direct measurement on a specially designed

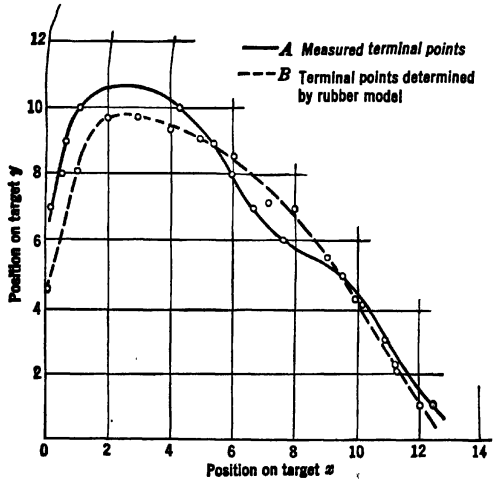


FIG. 12·14. Comparison of Location of Terminal Points of Electron Paths as Determined by Means of Experimental Tube and by the Rubber Model. (Courtesy *Proc. Inst. Radio Engrs.*, reference 10.)

tube with those found with the aid of the rubber model. The agreement is close enough to justify the use of the rubber model in most practical cases involving " two-dimensional " systems.

REFERENCES

1. F. S. WOODS, *Advanced Calculus*, Ginn and Company, Boston, 1926.
2. E. P. ADAMS, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, Smithsonian Institution, Washington, D.C., 1922.
3. O. SCHERZER, " On some aberrations of electron lenses," *Z. Physik*, Vol. 101, pp. 593–603, 1936.
4. R. GANS, " Electron paths in electron-optical systems," *Z. tech. Physik*, Vol. 18, pp. 41–48, February 1937.
5. A. RECKNAGEL, " On the theory of the electron mirror," *Z. Physik*, Vol. 104, pp. 381–394, 1936.
6. J. RAJCHMAN, " The dark current in electrostatic electron multipliers," *Arch. sci. phys. nat.*, Genève, Vol. 20, pp. 9–96, September–December 1938.
7. D. GABOR, " Mechanical tracer for electron trajectories," *Nature*, Vol. 139, p. 373, February 1937.
8. D. B. LANGMUIR, " Automatic plotting of electron trajectories," *Nature*, Vol. 139, p. 1066, June 1937.
9. P. H. J. A. KLEYNEN, " The motion of an electron in a two-dimensional electrostatic field," *Philips Tech. Rev.*, Vol. 2, pp. 321–352, 1937.
10. V. K. ZWORYKIN and J. A. RAJCHMAN, " The electrostatic multiplier," *Proc. Inst. Radio Engrs.*, Vol. 27, pp. 558–566, September 1939.

CHAPTER 13

GAUSSIAN DIOPTRICS OF ELECTROSTATIC LENSES

13·1. Axial Symmetry and Lens Action. The practical importance of electron optics follows in large part from the ability of electrostatic and magnetic fields to form electron images of objects irradiated by or emitting electrons. The condition which such fields must fulfil to possess this property is axial symmetry. It is quite possible to demonstrate the image-forming property of axially symmetric electrostatic fields with the aid of the ray equations; for example, Eqs. 12·9 and 12·10. However, such a demonstration would tend to obscure the extreme generality of the lens action of axially symmetric configurations. The proof given below, based on symmetry considerations only, shows, in effect, that any axially symmetric conservative force field[1] is capable of forming images. Similar considerations are found useful in examining the lens action of a magnetic field and, particularly, in deriving the general properties of the geometric aberrations.[2]

To return to the matter of immediate interest, any axially symmetric electrostatic field acts as an electron lens. It forms an electron image, real or virtual, of an electron-emissive or electron-transmitting object placed on the axis of symmetry. To a first approximation, this image is sharp and geometrically similar to the object. Normally, the properties of sharpness and geometrical similarity are confined to a small region of the image around the axis. Furthermore, they require that the electron rays participating in the image formation remain close to the axis and form small angles therewith.

To demonstrate the existence of this property of axially symmetric electrostatic fields, consider (Fig. 13·1) two planes normal to the axis of symmetry (z-axis): an object plane O with the coordinates x_o,y_o and, beyond the electrostatic fields, an image plane I (whose exact position is yet to be fixed) with the coordinates x_i,y_i. The origins of these two coordinate systems are assumed to lie on the z-axis and the pairs of x-axes and y-axes, respectively, to be coplanar. Imagine a third plane, the *aperture plane*, parallel to the object and image planes

[1] See section 10·3. The field must, in addition, be continuous about the axis of symmetry.

[2] See sections 15·3 and 16·2.

and intersecting the z-axis at A, with the coordinates x_a,y_a.[3] Then the path of any electron through the field will, in general, be completely defined if the coordinates of its intersection with the object plane,

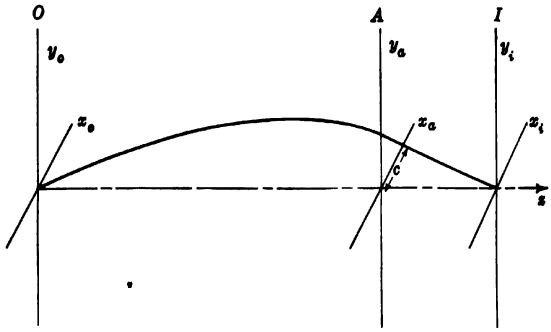


FIG. 13·1. Object, Aperture, and Image Planes.

x_o,y_o, and those with the aperture plane, x_a,y_a, are known. Accordingly its intersection x_i,y_i with the image plane is fully determined by x_o,y_o and x_a,y_a. Since the electrostatic force field will be assumed to be continuous throughout, it is thus possible to express the quantities x_i and y_i as power series of the variables x_o, y_o, x_a, and y_a:

$$x_i = a_o + a_1x_o + a_2y_o + a_3x_a + a_4y_a + a_5x_o^2 + a_6x_oy_o + \cdots \quad [13\cdot1a]$$

$$y_i = b_o + b_1x_o + b_2y_o + b_3x_a + b_4y_a + b_5x_o^2 + b_6x_oy_o + \cdots \quad [13\cdot1b]$$

The symmetry of the electrostatic field restricts the possible values of the coefficients in these relations so as to fulfil the conditions of the first-order imaging, enunciated above, for the field remains unaltered, and hence also the relations 13·1a and 13·1b, if (1) it is rotated through any angle ϕ about the axis, or if (2) it is reflected at any plane through the axis.

[3] For most purposes it is convenient to let the aperture plane coincide with the exit pupil of the lens system. The position of the exit pupil is found by continuing the tangents to the paraxial principal rays (that is, the central rays of the paraxial imaging pencils) at the image plane to their intersection with the optic axis. If an axially symmetric physical limiting aperture exists, the exit pupil is simply its image in image space. The aperture plane thus becomes a virtual, rather than a real, plane. Furthermore, the region between the aperture plane and image plane is "field free" by construction, since the "rays" are drawn as straight lines. It should be noted that all conclusions drawn in Chapters 13, 15, and 16 postulating a real aperture plane apply equally if the plane of the exit pupil is chosen as aperture plane.

Carrying out, first, a rotation through 180 degrees

$$\begin{array}{lll} x_i \to -x_i & x_o \to -x_o & x_a \to -x_a \\ y_i \to -y_i & y_o \to -y_o & y_a \to -y_a \end{array} \qquad [13\cdot2]$$

causes Eq. 13·1a to pass over into

$$-x_i = a_o - a_1 x_o - a_2 y_o - a_3 x_a - a_4 y_a + a_5 x_o^2 + a_6 x_o y_o + \cdots \qquad [13\cdot3]$$

If Eqs. 13·1a and 13·3 are added,

$$2a_o + 2a_5 x_o^2 + 2a_6 x_o y_o + \cdots = 0$$

This can be satisfied for all values of the coordinates only if the coefficients a_o, a_5, a_6, $\cdots$ of all even-power terms vanish individually. The same applies for b_o, b_5, b_6, $\cdots$.

Next, a rotation through 90°:

$$\begin{array}{lll} x_i \to y_i & x_o \to y_o & x_a \to y_a \\ y_i \to -x_i & y_o \to -x_o & y_a \to -x_a \end{array} \qquad [13\cdot4]$$

leads to

$$y_i = a_1 y_o - a_2 x_o + a_3 y_a - a_4 x_a + O(3) \qquad [13\cdot5a]$$

$$-x_i = b_1 y_o - b_2 x_o + b_3 y_a - b_4 x_a + O(3) \qquad [13\cdot5b]$$

Here $O(3)$ represents terms of the third and higher orders in the coordinates. Comparison with Eqs. 13·1a and 13·1b establishes the relations

$$b_1 = -a_2 \qquad b_2 = a_1 \qquad b_3 = -a_4 \qquad b_4 = a_3 \qquad [13\cdot6]$$

so that

$$x_i = a_1 x_o + a_2 y_o + a_3 x_a + a_4 y_a + O(3) \qquad [13\cdot7a]$$

$$y_i = -a_2 x_o + a_1 y_o - a_4 x_a + a_3 y_a + O(3) \qquad [13\cdot7b]$$

Finally, a reflection at the plane $x = 0$:

$$\begin{array}{lll} x_i \to -x_i & x_o \to -x_o & x_a \to -x_a \\ y_i \to y_i & y_o \to y_o & y_a \to y_a \end{array} \qquad [13\cdot8]$$

gives rise to

$$-x_i = -a_1 x_o + a_2 y_o - a_3 x_a + a_4 y_a + O(3) \qquad [13\cdot9a]$$

$$y_i = a_2 x_o + a_1 y_o + a_4 x_a + a_3 y_a + O(3) \qquad [13\cdot9b]$$

This is consistent with Eqs. 13·7 only if $a_2 = a_4 = 0$:

$$x_i = a_1 x_o + a_3 x_a + O(3) \qquad [13\cdot10a]$$

$$y_i = a_1 y_o + a_3 y_a + O(3) \qquad [13\cdot10b]$$

Consider now a ray leaving the origin in the object plane and intersecting the aperture plane in the point $x_a = c$, $y_a = 0$. For it,

$$x_i = a_3 c + O(3) \qquad [13\cdot11a]$$

$$y_i = 0 + O(3) \qquad [13\cdot11b]$$

As the position of the "image plane" has not been fixed, this means simply that the ray proceeds throughout in the xz-plane. Beyond the electric field it will be a straight line, inclined, in general, to the axis. If the ray converges toward the axis, the image plane will be considered to pass through the point of intersection of the ray with the z-axis; if it slopes away from the axis, a virtual image plane will be located by continuing the straight part of the ray path backwards to its intersection with the axis. In either case $x_i = 0$ for the ray considered; hence, quite generally, $a_3 = 0$, since $c \neq 0$, if the higher-order terms are negligible. To assure this, it is assumed that c is very small. Thus

$$x_i = a_1 x_o + O(3) \qquad [13\cdot12a]$$

$$y_i = a_1 y_o + O(3) \qquad [13\cdot12b]$$

Accordingly, if terms of the third and higher orders in the coordinates of the electron paths in the object and aperture planes are neglected, a sharp, geometrically similar image of the object is produced. This image may be either real or virtual. Its magnification, a_1, depends on the nature of the field; the sign of a_1 determines whether the image is erect or inverted. The sharp, geometrically faithful image obtained by putting the higher-order terms equal to zero is usually referred to as the first-order or Gaussian image. It is closely approximated by the real image in the neighborhood of the axis if imaging pencils of small aperture are employed.

In this chapter the higher-order terms will be neglected throughout. They will be considered in detail in Chapter 16, as they account for some of the major defects of the image, the geometrical aberrations.

13·2. Cardinal Points. Assume now that both the object and the image are in field-free space; that, in short, the focusing fields are confined to a limited region, indicated by the rectangle in Fig. 13·2. Let IB_i be the image of OB_o and assume that the rays $B_oF_iB_i$, incident parallel to the axis, and $B_oF_oB_i$, emerging parallel to the axis, are known. Their intersections with the axis, F_i and F_o, are designated as the image-side and object-side focal points of the lens system. They are the images formed by the lens of points on the axis, an infinite distance from the lens in object and image space, respectively. Thus all rays originally parallel to the axis converge at them after passing through the lens.

Furthermore, let the normal plane passing through the intersection of the continuation of B_iF_i with the continuation of the axis-parallel ray through B_o be denoted by H_i and that through the intersection of B_oF_o with the axis-parallel ray through B_i, by H_o. H_i and H_o are known as the image-side and object-side principal planes of the lens, respec-

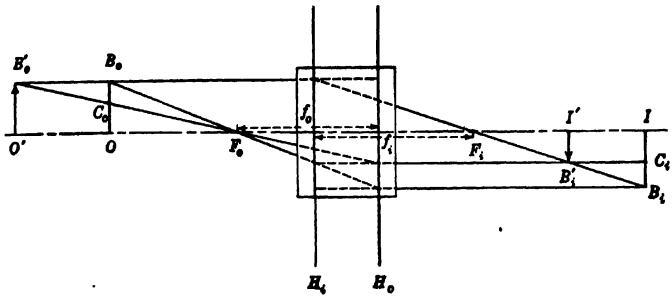


FIG. 13·2. Construction of the Image for Arbitrary Object Position.

tively. These planes have this property: The lens images them into each other with unity magnification. An electron-current (or intensity) distribution projected into the plane H_o from the object side will appear to exist in the plane H_i when viewed from the image side. Both object and image are here virtual.

The separations $F_oH_o = f_o$ and $H_iF_i = f_i$ are the object-side and image-side focal lengths of the lens. The geometrical construction in Fig. 13·2 shows that the magnification of the image is given by

$$a_1 = \frac{f_o}{F_oO} = -\frac{F_iI}{f_i} \qquad [13.13]$$

It is readily seen that, with the paths of the rays $B_oF_iB_i$ and $B_oF_oB_i$ given, both the position and the magnification of the image may be determined for any object position whatever, such as O'. If a ray is drawn from B_o' through F_o, it will intersect the original object plane in some point C_o and hence the image plane in a point C_i such that

$$IC_i = a_1 \cdot OC_o \qquad [13.14]$$

Furthermore, it will emerge from the system parallel to the axis, as the image of F_o is at infinity. Thus the image point B_i' — and hence the location and magnification of the image of $O'B_o'$ — are determined by the intersection of the parallel ray through C_i with the ray $B_o'B_oF_iB_i$. Again, for the new position of the object and image, the magnification

is seen to be given by

$$m = \frac{f_o}{F_o O'} = -\frac{F_i I'}{f_i} \qquad [13\cdot15]$$

If $O'H_o = u$, the *object distance*, and $H_i I' = v$, the *image distance*, the second part of Eq. 13·15 may be written

$$-\frac{f_o}{u - f_o} = -\frac{v - f_i}{f_i} \qquad \frac{f_i}{v} + \frac{f_o}{u} = 1 \qquad [13\cdot16]$$

It follows that

$$m = -\frac{f_o}{f_i}\frac{v}{u} \qquad [13\cdot17]$$

Equations 13·16 and 13·17 are the well-known lens equations. They permit the determination of the image distance (distance of the image from the image-side principal plane) and the image magnification from the object distance (distance of the object from the object-side principal plane) and the focal lengths of the lens. By knowing the four cardinal points, F_o, F_i, H_o, and H_i (focal points and principal planes), it is thus possible to construct the image for an object in an arbitrary position along the axis. It is necessary only that both object and image lie in field-free space, beyond the lens fields.

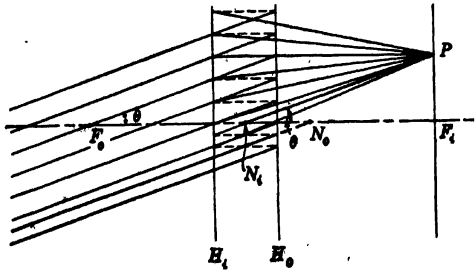


FIG. 13·3. Determination of the Nodal Points.

An additional pair of points characteristic of the system is the nodal points. They have this property: A ray in object space aimed at the object nodal point N_o has the same direction in image space, appearing to come from the image nodal point N_i. The determination of the position of these points is indicated on Fig. 13·3. Consider a parallel pencil of electron rays incident in object space at an angle to the axis θ. As this pencil may be regarded as diverging from an off-axis point at $-\infty$, it must focus in the image-space focal plane at the point P, a

distance $f_o \tan \theta$ from the axis, as follows from a consideration of the ray through the object-side focal point F_o.

One of the rays of the pencil will emerge from the lens at the same angle θ with respect to the axis. Designate the point of the axis at which it is aimed with N_o and the axial point from which it appears to emerge with N_i. Then, because of the parallelism of the incident and emergent ray and the fact that the planes H_o and H_i are imaged on each other with unity magnification — so that a ray aimed at a certain point on H_o appears to come from a point at an equal distance from the axis on H_i —

$$N_iN_o = H_iH_o \qquad [13\cdot18]$$

Furthermore, from the figure,

$$N_iF_i = f_o \qquad [13\cdot19]$$

so that, since $F_oF_i = f_o + f_i - H_iH_o$,

$$F_oN_o = f_i \qquad [13\cdot20]$$

Thus the nodal points are determined if the principal planes and focal points are known. The points H_o, H_i, F_o, F_i, N_o, and N_i constitute the cardinal points of a lens system. H_o and H_i are here used to denote both the principal planes and their intersection with the axis of symmetry.

13·3. Fundamental Theorems Regarding Electron Lenses. A simple relation between f_o and f_i can be deduced from the form of the paraxial ray equation (Eq. 12·10):

$$r'' + \frac{\Phi'}{2\Phi}r' + \frac{\Phi''}{4\Phi}r = 0 \qquad [13\cdot21]$$

If r_1 and r_2 are two independent solutions of this equation, there exists between them, according to the general theory of linear differential equations of the second order,[4] the relation

$$r_1r_2' - r_2r_1' = C\cdot\exp\left[-\int_{z_o}^{z}\frac{\Phi'}{2\Phi}dz\right] = C\cdot\left(\frac{\Phi_o}{\Phi}\right)^{1/4} \qquad [13\cdot22]$$

Assume that

$$r_1 = 1 \quad r_1' = 0 \quad \text{for} \quad z = z_o \qquad [13\cdot23a]$$

$$r_2 = 1 \quad r_2' = 0 \quad \text{for} \quad z = z_i \qquad [13\cdot23b]$$

z_o and z_i being points in object and image space, respectively. Then,

[4] See Adams, reference 1, Formula 8.410.

by the geometry of Fig. 13·4*a*,

$$r_1' = -\frac{1}{f_i} \qquad \text{for} \qquad z = z_i \tag{13·24a}$$

$$r_2' = \frac{1}{f_o} \qquad \text{for} \qquad z = z_o \tag{13·24b}$$

Substituting this in Eq. 13·22 leads to

$$\frac{1}{f_o} = C \qquad \frac{1}{f_i} = C\left(\frac{\Phi_o}{\Phi_i}\right)^{1/2}$$

or

$$\frac{f_o}{f_i} = \left(\frac{\Phi_o}{\Phi_i}\right)^{1/2} \tag{13·25}$$

Thus the object-side and image-side focal lengths are to each other as the square root of the potential in the object and the image space.

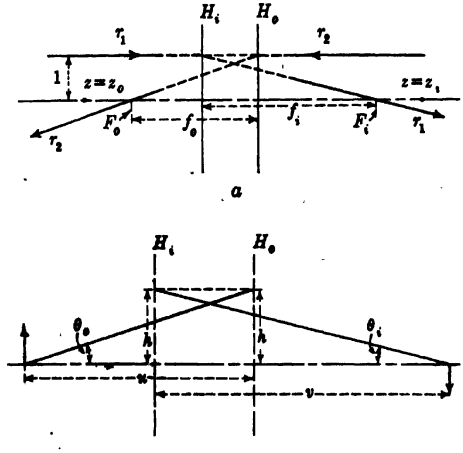


FIG. 13·4. Relation between (*a*) the Focal Lengths and (*b*) the Angular Divergence in Object and Image Space.

In other words, they are proportional, as in light optics, to the respective indices of refraction. In particular, for a *unipotential lens*, for which the potentials in object and image space are the same, the two focal lengths become identical:

$$f_o = f_i = f \tag{13·26a}$$

As a consequence, furthermore, the nodal points of such lenses fall into the corresponding principal planes:

$$N_o \equiv H_o \qquad \text{and} \qquad N_i \equiv H_i \tag{13·26b}$$

An important relation exists also between the relative divergence of the imaging pencils in image and object space and the magnification of the image. Thus let the angle of inclination of a ray leaving the axial object point be θ_o and the inclination of the same ray in image space be θ_i. Then, in view of the definition of the principal planes, the height of incidence of a ray in object space on H_o must be the same as the apparent height of incidence of the ray in image space on H_i. Thus, let these heights be equal to h (Fig. 13·4b). The two angles are then given by

$$\theta_o = \frac{h}{u} \qquad \text{and} \qquad \theta_i = -\frac{h}{v}$$

Accordingly

$$\frac{\theta_o}{\theta_i} = -\frac{v}{u}$$

By Eqs. 13·17 and 13·25 this leads to the relation

$$m = -\frac{\theta_o}{\theta_i}\left(\frac{\Phi_o}{\Phi_i}\right)^{1/2} \tag{13·27}$$

This is a special form, applicable for paraxial rays in centered optical systems, of the very general theorem of Clausius or Helmholtz-Lagrange, which has been discussed in section 10·5.

It may have been noted that in Figs. 13·2, 13·3, and 13·4 the principal planes and nodal points have been shown " crossed." Thus the object-side principal plane H_o and the nodal point N_o were drawn closer to image space (F_i) than the image-side principal plane H_i and the nodal point N_i, respectively. This may be shown to apply quite generally for electron lenses.

Consider (Fig. 13·5) a ray in object space aimed at the nodal point N_o. The corresponding curve $R(z) = r(z)\cdot\Phi^{1/4}(z)$ will, in object space, be represented by a straight line inclined, in general, at a different angle to the axis, but aimed at the same point. The curvature of $R(z)$ is given, for paraxial rays, by

$$R'' = -TR \tag{13·28}$$

where $T = \frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2$ is always positive. Hence the curve $R(z)$ is always concave toward the axis. Accordingly, $R(z)$ will intersect the

axis to the left of N_o, for example, at the point P. Consider now the tangent to the curve $R(z)$ at any point Q_1, corresponding to the abscissa

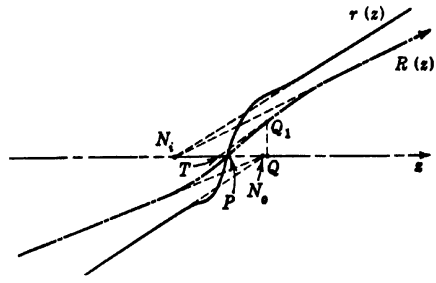


FIG. 13·5. The Crossing of the Nodal Points.

Q, beyond P. The slope decreases monotonically from P onward, so that the distance from Q of the intersection T of the tangent with the axis is

$$TQ = \int_P^Q \frac{R'\,dz}{R'(Q)} > PQ \qquad [13\cdot29]$$

Therefore T and, also, the point of intersection N_i of the asymptotic

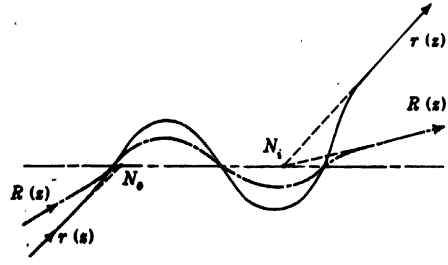


FIG. 13·6. Possible Location of the Nodal Points for a System Forming Intermediate Images.

tangent to $R(z)$ in image space with the axis lie to the left of P and, hence, of N_o. Thus the nodal points and, hence, the principal planes, are, in fact, crossed.

This proof does not apply when intermediate images are formed. In this case (Fig. 13·6) the ray aimed at N_o may intersect the axis several times. The compound microscope considered as a whole is a lens system of this type.

Another important restrictive property of electron lenses bounded by regions of constant potential (object and image space) may be deduced by utilizing the variable R. This is the nonexistence of negative lenses,

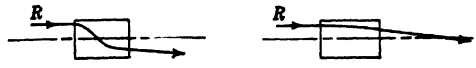


FIG. 13·7. Nonexistence of Negative Electron Lenses.

that is, lenses for which rays incident parallel to the axis emerge diverging from the axis, without, in the interim, having crossed the axis. As r is represented in object space by a line parallel to the axis, $R = r\Phi^{1/4}$ is represented by a similar parallel line in this region. The curvature of of R is, however, always toward the axis. There are, therefore, only two possibilities (Fig. 13·7):

a. R, and hence r, crosses the axis within the lens.

b. R does not cross the axis within the lens, but emerges, in image space, as a straight line converging toward the axis. Then $r = R\Phi^{-1/4}$ is also represented by a straight line converging toward the axis in image space, as $\Phi^{-1/4}$ is here a positive constant.

In neither case does the electron lens act as a negative lens.

13·4. Classification of Electron Lenses. Electrostatic electron lenses can be divided conveniently into the following six classes:

A. Screen Lenses. The most exact analogue of the light-optical glass lens is, without doubt, the screen lens, studied in some detail by Knoll and Ruska[5] in 1932. Here two parallel wire screens at different potentials simulate the glass surface of a light lens, dividing space into two equipotential regions, inside and outside the lens (Fig. 13·8). Numerous other electron lenses using wire screens to form suitable equipotential surfaces have been suggested. In all cases the image quality is impaired by the inevitable nonuniformity of field and potential in the plane of the wire screen. For this reason screen lenses have not attained practical importance. The replacement of the wire screens by thin metallic foil transparent to electrons, which would overcome this difficulty, has not been attempted because of the delicacy of such foils. Since the electrostatic field is discontinuous in screen lenses and foil

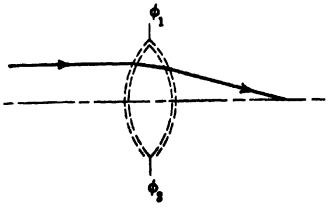


FIG. 13·8. A Screen Lens.

[5] See reference 2.

lenses, the general theorems deduced in the preceding section do not apply to them.

B. Unipotential or Unit Lenses, Characterized by Equal Constant Potentials in Object and Image Space. Among the typical electron lenses, that is, those in which the region traversed by the electrons between object and image is free from matter, the unipotential lenses resemble most closely the commonly employed types of lenses in light optics — glass lenses and lens systems in air. As indicated above (Eq. 13·26), the image-side and object-side focal lengths are here equal and the corresponding principal and nodal points coincide.

C. Bipotential or Immersion Lenses, Characterized by Unequal Constant Potentials in Object and Image Space. The second designation is derived from their analogy to the oil-immersion objectives of the light microscope, for which object and image are placed in media of different index of refraction, that is, oil and air, respectively. The corresponding principal and nodal points are separated by a distance equal to the difference of the image- and object-side focal lengths:

$$H_oN_o = H_iN_i = f_i - f_o$$

The image-side and object-side focal lengths, themselves, are to each other as the square roots of the potentials in image space and object space, respectively (Eq. 13·25).

D. Single-Aperture Lenses. These comprise the lens fields about an aperture in an electrode which separates two regions of different, constant electric field. Since, here, the electron paths in the object and image fields are not straight lines, but parabolas, the considerations of the last two sections, concerning cardinal points, do not apply. Similarly, the lens equations (13·16 and 13·17) cannot be utilized. It will be seen, however, that the image magnification and image distance can be found from the position of the focal points and principal planes, if the points and planes are redefined in a suitable manner.

E. Cathode Lenses. These are lens fields which are terminated on one side by a surface at zero potential (cathode) normal to the optic axis. The cathode constitutes at the same time the object surface, so that object space may be regarded as contracted into a surface at zero potential. Cathode lenses occur in any system in which a cathode is imaged by its own electron emission, be it thermionic, photoelectric, secondary, or field emission.[6]

F. Space-Charge Lenses. All the above types of lenses may be modified in their action by the presence of space charge — in the form

[6] The common designation of cathode lenses as *immersion objectives* is, in the exact sense, a misnomer.

either of large concentrations of electrons or of positive ions. Under certain circumstances these may exercise a strong, independent lens action. Thus the well-known experiments with *thread beams*[7] (narrow, luminous electron beams in an imperfectly evacuated vessel held together by the attraction of the positive-ion space charge formed along the beam) rely on this effect. Positive-ion concentrations in gas discharges have been employed by von Borries and Ruska[8] for forming electron images of wire screen objects. However, space-charge lenses have been found to be too difficult to control to render them valuable in the formation of images. Accordingly, space charge will here be considered only as a disturbing factor in image formation.[9]

The classification of electron lenses here given is, of course, by no means unique. An alternative, which cuts across this grouping, divides all electron lenses into long and short (or thin) lenses. Thin lenses respond to a particularly simple treatment and will hence be considered first.

13·5. The Thin Lens. A unipotential or an immersion lens is considered "thin" or "short" if the refracting field is short compared to its focal length. In this case it is permissible to regard the total lens action as taking place at a plane intersecting the axis in $H_o \equiv H_i \equiv H$. The nodal point $N_o \equiv N_i \equiv N$ lies at a distance $f_i - f_o$ from this principal, or lens, plane. If u is the object distance from the lens plane and v the image distance, they are related, according to Eqs. 13·16 and 13·25, by

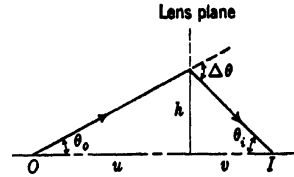


FIG. 13·9. Refraction of a Ray at the Lens Plane.

$$\frac{\Phi_o^{1/2}}{u} + \frac{\Phi_i^{1/2}}{v} = \frac{\Phi_o^{1/2}}{f_o} = \frac{\Phi_i^{1/2}}{f_i} \qquad [13\cdot30]$$

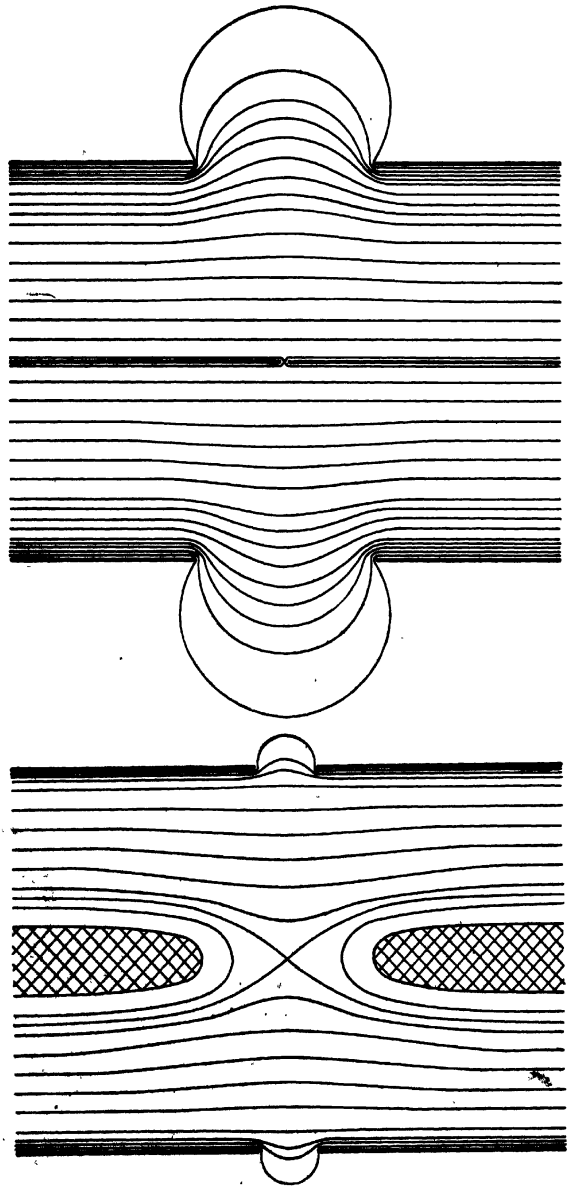
If θ_o is the angle of inclination of a ray from the axial point O in object space and θ_i that of the corresponding ray in image space (Fig. 13·9),

$$\Phi_o^{1/2}\theta_o + \Phi_i^{1/2}\theta_i = \frac{\Phi_o^{1/2}h}{f_o} \qquad [13\cdot31]$$

In particular, if $\Phi_o = \Phi_i$, that is, in case of a unit or unipotential lens,

$$\Delta\theta = \theta_o + \theta_i = \frac{h}{f} = \text{const}\cdot h \qquad [13\cdot32]$$

[7] See Brüche, reference 3.
[8] See reference 4.
[9] See section 16·9.



(a). Unipotential Lens with Large Central Aperture. (b). Unipotential Lens with Small Central Aperture

FIG. 13·10. Unipotential Lenses Formed by Three Apertures.

Thus a thin lens between two media of equal potential (equal index of refraction) alters the angle of inclination of any incident ray by an amount proportional to the height of incidence, the factor of proportionality being the refractive power of the lens, $1/f$.

If the refractive index on the two sides of the lens differs, the factor of proportionality becomes $1/u + 1/v$. This is a constant for a fixed position of the intersection of the ray with the axis in object space, but varies as this point is displaced.

It is relatively easy to obtain an expression for the focal length of a thin lens, since, here, the variable $R = r\Phi^{1/4}$ may be regarded as constant within the lens field.[10] Thus an integration of the ray equation (13·28)

$$R'' = -\frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 R$$

leads to

$$R'(b) - R'(a) = -\int_a^b \frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 R\, dz \qquad [13\cdot33]$$

where a and b signify points just to the left and right of the lens, respectively. For a ray incident parallel to the axis,

$$R'(a) = 0 \qquad \text{and} \qquad \frac{1}{f_i} = -\frac{R'(b)}{R(a)}\left(\frac{\Phi_o}{\Phi_i}\right)^{1/4}$$

Hence, since R is assumed to be constant ($= R(a)$) between a and b,

$$\frac{1}{f_i} = \frac{3}{16}\left(\frac{\Phi_o}{\Phi_i}\right)^{1/4}\int_a^b \left(\frac{\Phi'}{\Phi}\right)^2 dz \qquad [13\cdot34]$$

Similarly,

$$\frac{1}{f_o} = \frac{3}{16}\left(\frac{\Phi_i}{\Phi_o}\right)^{1/4}\int_a^b \left(\frac{\Phi'}{\Phi}\right)^2 dz \qquad [13\cdot35]$$

13·6. Unipotential Lenses. The simplest and most thoroughly studied electrostatic unipotential lens is the symmetrical three-aperture lens. Figures 13·10*a* and 13·10*b* show the potential distributions in two lenses of this type. The first has a large, the second a small central

[10] This assumption leads to a smaller error than setting r = const. As has been seen, R is uniformly concave to the axis, irrespective of the character of the lens field; r, however, varies so that its value is larger in the converging parts of the lens than in the diverging parts of it. Consequently, putting r = const invariably leads to too low a value for the refractive power $1/f$.

aperture. The succeeding figures, Figs. 13·11 and 13·12, indicate the lens properties of these same lenses as function of the ratio of the voltages applied to the central electrode (V_c) and the outer electrodes (V_o).

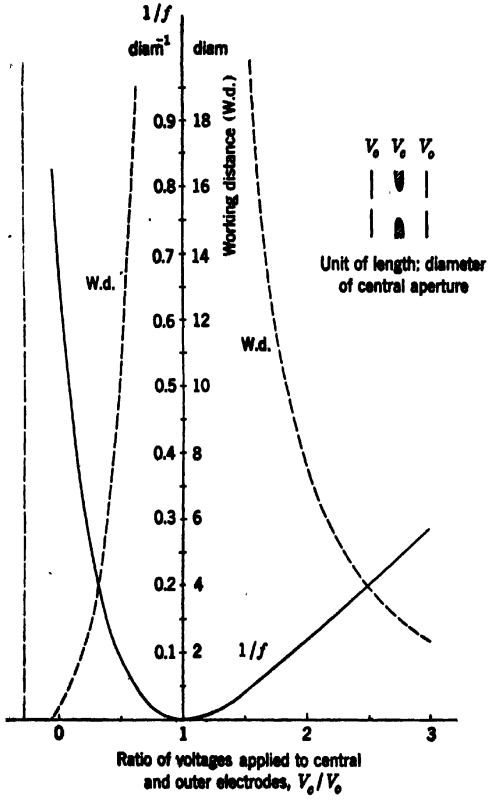


Fig. 13·11. Variation with the Ratio of the Applied Voltages of the Refractive Power and the "Working Distance" for the Electrostatic Lens Shown in Fig. 13·10*a*.

The characteristic quantities plotted are the *refractive power*, or reciprocal focal length, $1/f$, and the *working distance*, or separation of the focal point from the plane of the nearest outer aperture. The latter notation is derived from the fact that, for large image magnifications, the object position and the focal point coincide approximately.

A cursory examination of the refractive power curves indicates that for a given electron velocity strong lenses can be obtained much more

readily by making the central electrode negative with respect to the outer electrodes than by making it positive. Consequently, lenses of this type are normally operated with the central electrode at a potential

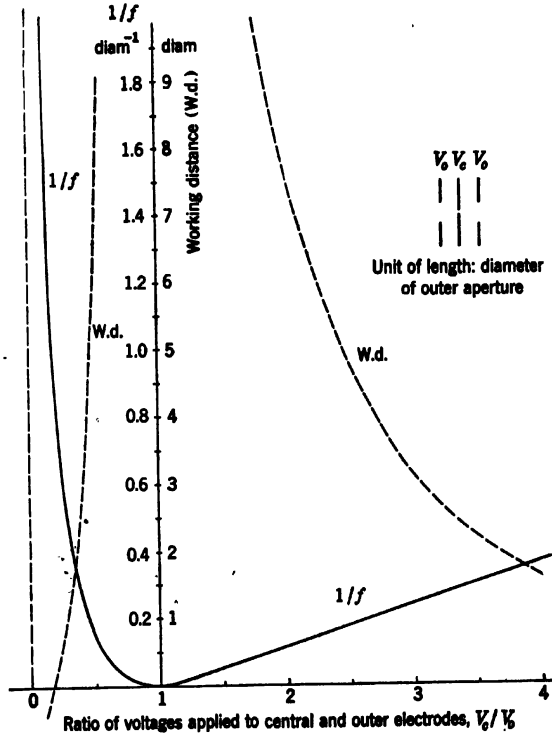


FIG. 13·12. Variation with the Ratio of the Applied Voltages of the Refractive Power and the "Working Distance" for the Electrostatic Lens Shown in Fig. 13·10*b*.

which is low compared to that of the outer electrodes. The lowest potential of the central electrode for which such fields will act as lenses is that at which the axial potential at the center of the lens is just reduced to zero. For the large-aperture lens this occurs at $V_c = -0.278V_o$; for the small-aperture lens at $V_c = -0.013V_o$. A practical limit — for most types of objects — is reached much earlier. The working distance should be positive and, preferably, large enough so that the object itself does not modify the lens field. This is fulfilled if the working distance is equal to the diameter of the outer aperture. The condi-

tion is satisfied for the large-aperture lens for $V_c > 0$; for the small-aperture lens for $V_c > 0.25V_o$. The relatively small outer apertures in the large-aperture lens serve the purpose of reducing the least permissible value of the working distance.

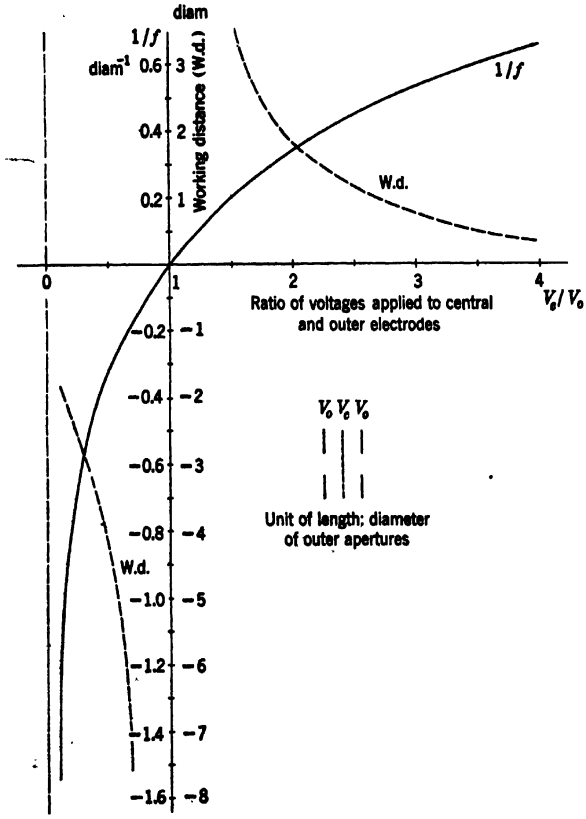


FIG. 13·13. Variation with the Ratio of the Applied Voltages of the Refractive Power and the "Working Distance" for the Lens in Fig. 13·10*b* with the Central Aperture Replaced by a Conducting Membrane.

Both the lenses just discussed are of course positive irrespective of whether the potential of the central electrode is positive or negative with respect to that of the outer electrodes. This condition no longer holds if the central electrode is replaced by a fine metal screen or an electron-

permeable conducting membrane.[11] The resulting *screen lens* is positive if the central electrode is made positive with respect to the outer electrodes; negative, if it is made negative with respect to them. However, lower positive voltages applied to the central electrode now suffice to attain a given refractive power, since there is no divergent lens action near the center of the lens. This is evident from the curves in Fig. 13·13. The refractive power is now a monotonically increasing function of the voltage ratio with uniformly decreasing slope. It is worth noting that in this case the principal planes are not crossed; the separation of the focal points from the center of the lens is greater than the focal length of the lens. If a metal screen is employed as central electrode, a projection of the screen will, in general, appear superposed on any image formed by the lens. Only in the very special case when the lens is made so strong that initially parallel ray pencils converge in the central plane of the lens will this effect be entirely absent. Under such circumstances, however, the utility of the lens is greatly restricted.

13·7. Single-Aperture Lenses. In some respects the single circular aperture in a plane electrode separating two regions of different field forms the most elementary electrostatic lens. However, the fact that it is now necessary to define object space and image space as regions of constant electric field rather than as regions of constant potential demands a treatment which deviates from that accorded systems bounded by field-free regions. If the plane of the aperture has the abscissa $z = 0$, the potential of the aperture electrode is Φ_o, and the potential gradient at large distances from the aperture, to the right and to the left, is Φ_2' and Φ_1', respectively, the electron paths in object and image space are parabolas given by

$$r = k_1\left[\frac{\Phi_o + \Phi_1' z}{|\Phi_1'|}\right]^{1/4} + k_1' \qquad [13\cdot36a]$$

and

$$r = k_2\left[\frac{\Phi_o + \Phi_2' z}{|\Phi_2'|}\right]^{1/4} + k_2' \qquad [13\cdot36b]$$

Here k_1, k_2, k_1', and k_2' are constants determined by the initial radial separation and slope of the electron path considered. $\Phi_o + \Phi_1' z$ is the potential in object space, $\Phi_o + \Phi_2' z$ that in image space. Assume now that the rays incident on the lens field parallel to the axis intersect the axis at points lying in regions where the potentials are adequately rep-

[11] See Knoll and Weichardt, reference 5

resented by these quantities. Furthermore (Fig. 13·14), define these points as the "focal points" F_i' and F_o' with the abscissas z_i' and z_o',

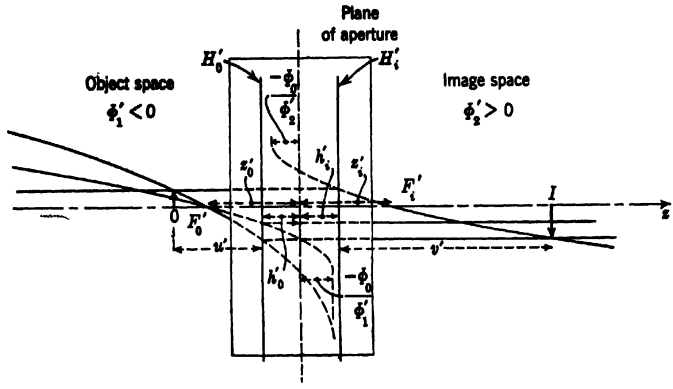


FIG. 13·14. Cardinal Points and Image Construction for a Single-Aperture Lens.

the *primes* indicating an unconventional definition. The paths of rays incident parallel to the axis at a distance r_o and r_i from it are given by

$$r = r_o \frac{\left[\frac{\Phi_o + \Phi_2' z_i'}{|\Phi_2'|}\right]^{1/2} - \left[\frac{\Phi_o + \Phi_2' z}{|\Phi_2'|}\right]^{1/2}}{\left[\frac{\Phi_o + \Phi_2' z_i'}{|\Phi_2'|}\right]^{1/2} - \left[\frac{\Phi_o + \Phi_2' h_i'}{|\Phi_2'|}\right]^{1/2}} \qquad [13\cdot37a]$$

and

$$r = r_i \frac{\left[\frac{\Phi_o + \Phi_1' z_o'}{|\Phi_1'|}\right]^{1/2} - \left[\frac{\Phi_o + \Phi_1' z}{|\Phi_1'|}\right]^{1/2}}{\left[\frac{\Phi_o + \Phi_1' z_o'}{|\Phi_1'|}\right]^{1/2} - \left[\frac{\Phi_o + \Phi_1' h_o'}{|\Phi_1'|}\right]^{1/2}} \qquad [13\cdot37b]$$

respectively. The *principal planes* H_o' and H_i', with the abscissas $z = h_o'$ and h_i', respectively, are the planes normal to the axis which contain the points of intersection of the parabolic paths (Eqs. 13·37) with the original axis-parallel paths. Then, if the image distance v' is the distance of the image from H_i' and if u' is the distance of the object from H_o', the magnification m and position of the image can be determined if the newly defined cardinal points and the object distance are known:

$$m = -\frac{(\Phi_o + \Phi_1' z_o')^{1/2} - (\Phi_o + \Phi_1' h_o')^{1/2}}{(\Phi_o + \Phi_1' z_o')^{1/2} - (\Phi_o + \Phi_1' [u' + h_o'])^{1/2}} \qquad [13\cdot38]$$

$$\frac{(\Phi_o + \Phi_2'[v' + h_i'])^{1/2} - (\Phi_o + \Phi_2' z_i')^{1/2}}{(\Phi_o + \Phi_2' z_i')^{1/2} - (\Phi_o + \Phi_2' h_i')^{1/2}} = \frac{(\Phi_o + \Phi_1' z_o')^{1/2} - (\Phi_o + \Phi_1' h_o')^{1/2}}{(\Phi_o + \Phi_1'[u' + h_o'])^{1/2} - (\Phi_o + \Phi_1' z_o')^{1/2}} \quad [13\cdot39]$$

The lens equations in terms of these newly defined lens parameters are thus considerably more complex than the equations for an ordinary lens, bounded by regions of constant potential or refractive index. They are,

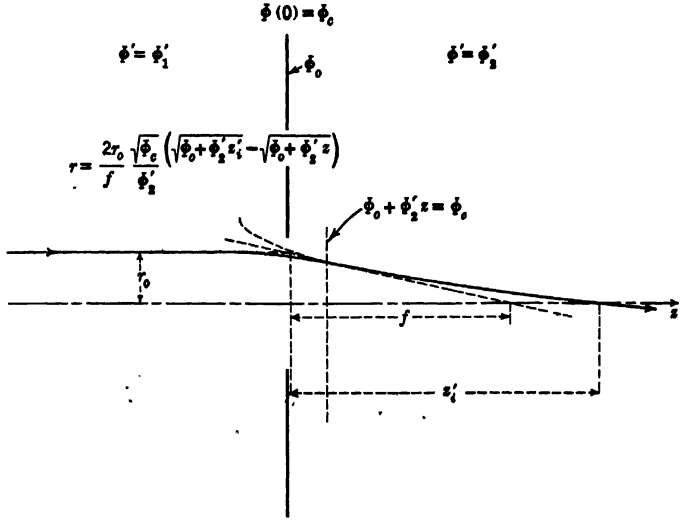


FIG. 13·15. Definition of the Focal Length f of a Single-Aperture Lens.

accordingly, rarely used. Normally, the focal length of a single-aperture lens is not defined by $z_i' - h_i'$, but as the axial distance between the intersections of the tangent to the parabolic path Eq. 13·37a at $\Phi_o + \Phi_2' z = \Phi_c$ $\left(\Phi_c = \dfrac{D}{2\pi}(\Phi_2' - \Phi_1') + \Phi_o$ is the potential at the center of an aperture of diameter $D\right)$ with the parallel incident ray and with the axis (Fig. 13·15). In the special case (Fig. 13·18a) in which the field in image space is equal to zero, the two definitions of the focal length are identical. The refractive power $1/f$, f being defined in accord with Fig. 13·15, indicates the degree to which a ray is bent by the inhomogeneity of the field about the aperture (Eq. 13·32).

If the fields on either side of the aperture are small compared to the ratio of the aperture potential to the aperture diameter, formulas for this refractive power can be derived very simply. In this case the potential within the region of changing field may be set equal to Φ_c, the potential at the center of the aperture. Consider a cylinder, symmetric

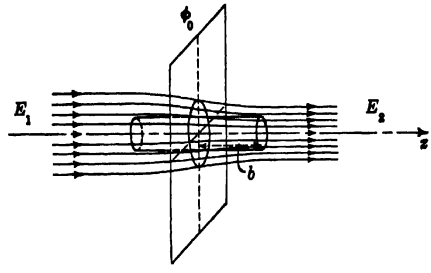


FIG. 13·16. Application of Gauss' Theorem to the Field Near an Aperture Separating Two Regions of Constant Field.

about the axis and the plane of the aperture, of length $2b$, just long enough so that the fields across the ends of the cylinder correspond, substantially, to the uniform fields $E_1 = -\Phi_1'$ and $E_2 = -\Phi_2'$, prevailing at large distances from the aperture (Fig. 13·16). As the cylinder contains no charge, the integral of the normal component of the field over its surface must vanish. Thus, by Gauss' theorem,

$$2\pi r \int_{-b}^{b} E_r \, dz = -\pi r^2 (E_2 - E_1) \qquad [13\cdot40]$$

where E_r is the radial component of the electric field across the mantel surface of the cylinder. In view of the assumptions made with regard to the field strengths on the two sides of the aperture, a paraxial electron incident parallel to the axis will change neither its separation from the axis nor its axial velocity to an appreciable extent while passing through the cylinder. The axial velocity is given throughout by

$$v_z = \left(\frac{2e\Phi_c}{m}\right)^{1/2} \qquad [13\cdot41]$$

However, if the separation of the electron from the axis is r, it is given a radial component of momentum equal to

$$mv_r = -\int eE_r \, dt = -\frac{\int_{-b}^{b} eE_r \, dz}{v_z} \qquad [13\cdot42]$$

It will hence emerge from the cylinder at an angle α relative to the axis

$$\tan \alpha = \frac{v_r}{v_z} = \frac{r(E_2 - E_1)}{4\Phi_c} \qquad [13\cdot43]$$

The integration of Eq. 13·42 is made possible by the fact that, to a first approximation, $E_r = (r/2)\Phi'' = -(r/2)\, dE_z/dz$. The focal length of the single-aperture lens thus becomes

$$f = -\frac{r}{\tan \alpha} = \frac{4\Phi_c}{E_1 - E_2} = \frac{4\Phi_c}{\Phi_2' - \Phi_1'} \qquad [13\cdot44]$$

This result[12] is, of course, only an approximation. It can be obtained much more simply by an application of Gans' method of determining electron paths.[13] For this purpose the potential is represented by a broken line with a single break point at $z = 0$ (Fig. 13·17). Integration of the ray equation

$$r'' = -\left(\frac{\Phi'}{2\Phi}\right) r' - \left(\frac{\Phi''}{4\Phi}\right) r$$

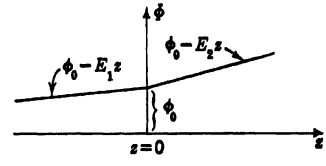


FIG. 13·17. Application of Gans' Method for Deriving the Lens Action of an Aperture.

over an infinitesimal segment including the point $z = 0$ yields, as Φ and r may be regarded as constant and r' and Φ' are finite,

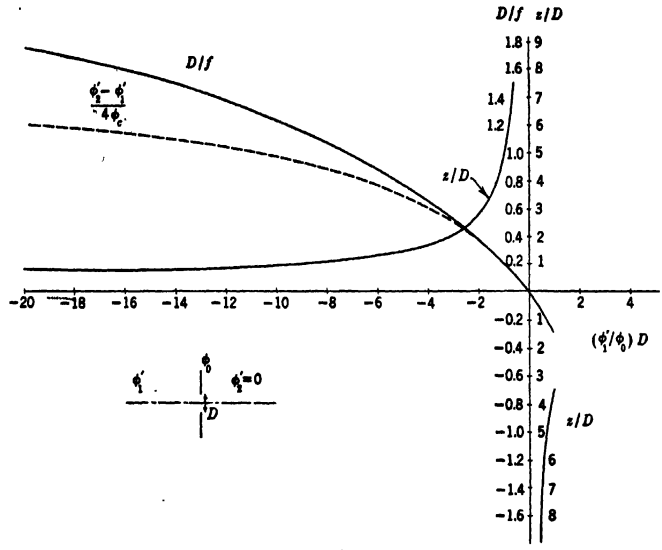
$$r_2' - r_1' = -\frac{(\Phi_2' - \Phi_1')r}{4\Phi_c} = \frac{r(E_2 - E_1)}{4\Phi_c} \qquad [13\cdot45]$$

With $r_1' = 0$ and $r_2' = \tan \alpha$ this becomes identical with Eq. 13·43.

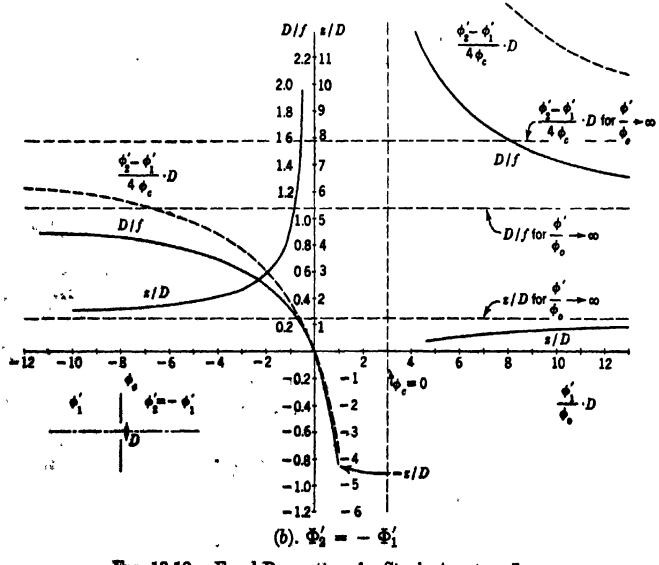
Figures 13·18*a* and *b* compare the refractive power of single-aperture lenses as determined by numerical integration of the ray equation with that given by the approximate formula (13·44). The first graph (Fig. 13·18*a*) refers to the case $\Phi_2' = 0$. The diameter of the aperture, D, serves as unit of length. The abscissas are the ratio of the potential gradient to the left of the aperture, Φ_1', to the potential of the apertured electrode, Φ_0. The distance z of the focal point in field-free image space from the plane of the aperture is also plotted in this figure. This is found to be greater than the focal length f. The principal planes are

[12] See Davisson and Calbick, reference 6.

[13] See section 12·4.



(a). $\Phi_2' = 0$.



(b). $\Phi_2' = -\Phi_1'$

FIG. 13-18. Focal Properties of a Single-Aperture Lens.

not crossed. It is seen that, as the field to the left of the aperture is made more strongly retarding, the refractive power given by Eq. 13·44 lags behind the true refractive power. On the other hand, for accelerating fields Φ_1' Eq. 13·44 yields a very slightly too high (negative) refractive power.

Figure 13·18*b* covers the case of equal fields of opposite sign on the two sides of the aperture ($\Phi_2' = -\Phi_1'$). Here the refractive power $1/f$, f being defined by Fig. 13·15, is compared with that given by Eq. 13·44. The values given by the approximate formula Eq. 13·44 are here seen to be high for convergent lens action, low for divergent lens action. The curves in the upper right quadrant refer to the case when both Φ_1' and Φ_o are negative, while Φ_c, the potential at the center of the aperture, is still positive.

In general, Eq. 13·44 is seen to give a good approximation of the refractive power of a single-aperture lens, provided that the fields involved are weak compared to the ratio of the potential of the aperture electrode to the aperture diameter. This is to be expected from its derivation. Single-aperture lenses are not subject to the restrictive properties applying to the more typical unipotential and immersion lenses (section 13·3).

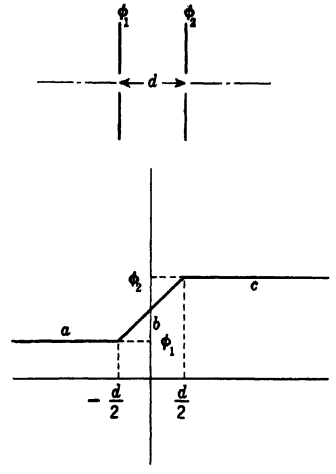


FIG. 13·19. Approximation of the Potential Distribution between Two Apertures for the Application of Gans' Method.

13·8. Immersion Lenses. Possibly the simplest immersion lens consists of two apertured electrodes at potentials Φ_1 and Φ_2, separated by a distance d, object and image space having constant potentials Φ_1 and Φ_2, respectively (Fig 13·19). This case may also be treated by Gans' method. For this purpose the lens field is divided into three sections:

I. $\Phi = \Phi_1$ for $z < -d/2$.

II. $\Phi = (\Phi_2 + \Phi_1)/2 + (\Phi_2 - \Phi_1)z/d$ for $-d/2 < z < d/2$.

III. $\Phi = \Phi_2$ for $x > d/2$.

In the terminology adopted in section 12·4, the following relations are

obtained successively for a ray incident parallel to the axis at a height r_o:

$$r'_{ab+} = -\frac{r_o}{d}\frac{\Phi_2 - \Phi_1}{4\Phi_1} \qquad C_b = -\frac{r_o}{d}\frac{\Phi_2 - \Phi_1}{4\Phi_1^{1/2}}$$

$$r_{bc} = r_o\left(1 - \frac{\Phi_2^{1/2} - \Phi_1^{1/2}}{2\Phi_1^{1/2}}\right) = r_o\frac{3\Phi_1^{1/2} - \Phi_2^{1/2}}{2\Phi_1^{1/2}}$$

$$r'_{bc-} = -\frac{r_o}{d}\frac{\Phi_2 - \Phi_1}{4(\Phi_2\Phi_1)^{1/2}}$$

$$r'_c = -\frac{r_o}{d}\frac{\Phi_2 - \Phi_1}{4(\Phi_2\Phi_1)^{1/2}} + \frac{r_o(3\Phi_1^{1/2} - \Phi_2^{1/2})}{8\Phi_2\Phi_1^{1/2}}\frac{\Phi_2 - \Phi_1}{d}$$

$$= -\frac{3r_o}{8d\Phi_2\Phi_1^{1/2}}(\Phi_2^{1/2} - \Phi_1^{1/2})(\Phi_2 - \Phi_1)$$

$$f_2 = -\frac{r_o}{r'_c} = \frac{8d}{3\left(\left[\frac{\Phi_2}{\Phi_1}\right]^{1/2} - 1\right)\left(1 - \frac{\Phi_1}{\Phi_2}\right)} \qquad [13\cdot46a]$$

$$f_1 = \left(\frac{\Phi_1}{\Phi_2}\right)^{1/2} f_2 = \frac{8d}{3\left(1 - \left[\frac{\Phi_1}{\Phi_2}\right]^{1/2}\right)\left(\frac{\Phi_2}{\Phi_1} - 1\right)} \qquad [13\cdot46b]$$

The principal planes are located at $z = h_1$ and $z = h_2$, respectively:

$$h_2 = \frac{d}{2} + \frac{r_o - r_{bc}}{r'_c} = \frac{d}{2} - \frac{4d\Phi_2}{3(\Phi_2 - \Phi_1)} \qquad [13\cdot47a]$$

$$h_1 = -\frac{d}{2} - \frac{4d\Phi_1}{3(\Phi_2 - \Phi_1)} \qquad [13\cdot47b]$$

Accordingly, the separation between principal planes is

$$h_2 - h_1 = -\frac{d}{3} \qquad [13\cdot48]$$

Figure 13·20 compares the values of the refractive power as given by Eq. 13·46*a* with those determined by numerical integration for the specific case in which the separation of the aperture planes is twice as large as the aperture diameters. It is seen that under these circumstances the re ractive powers determined by Eq. 13·46*a* are too large, the difference being especially great if the ratios of the applied potentials are large. The positions of the focal points (z_2) are, correspondingly,

farther removed from the plane of symmetry than is indicated by Eqs. 13·46*a* and 13·47*a*. However, it is to be expected that these differences will approach zero as the apertures in the electrodes are made smaller and smaller, other quantities remaining unchanged.

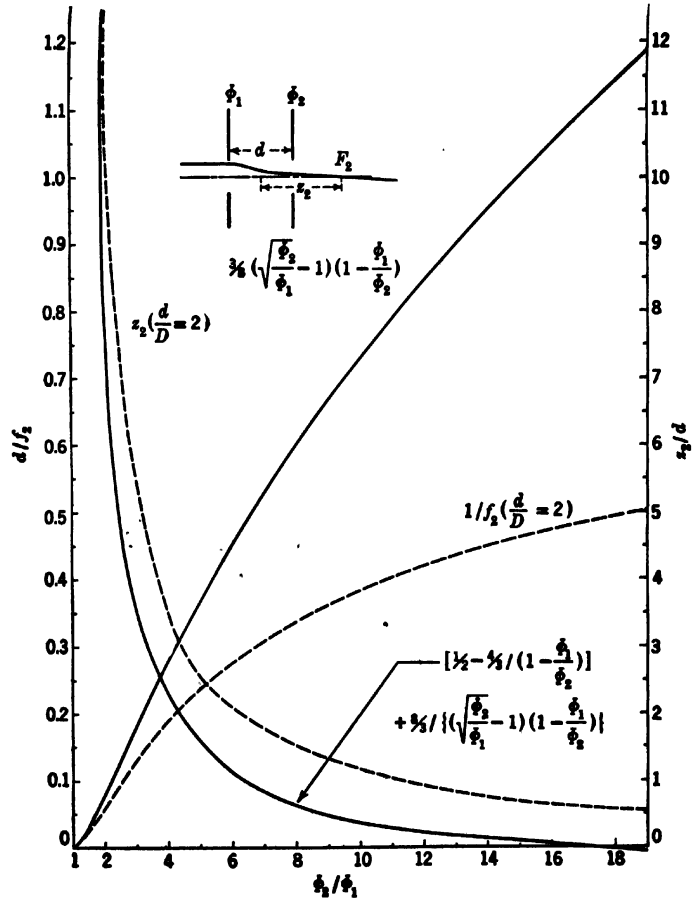


FIG. 13·20. Focal Properties of the Two-Aperture Lens.

A second type of immersion lens, which finds very wide application, is formed by two coaxial cylinders at different potentials. If they have equal diameters and are separated by a very small distance, the po-

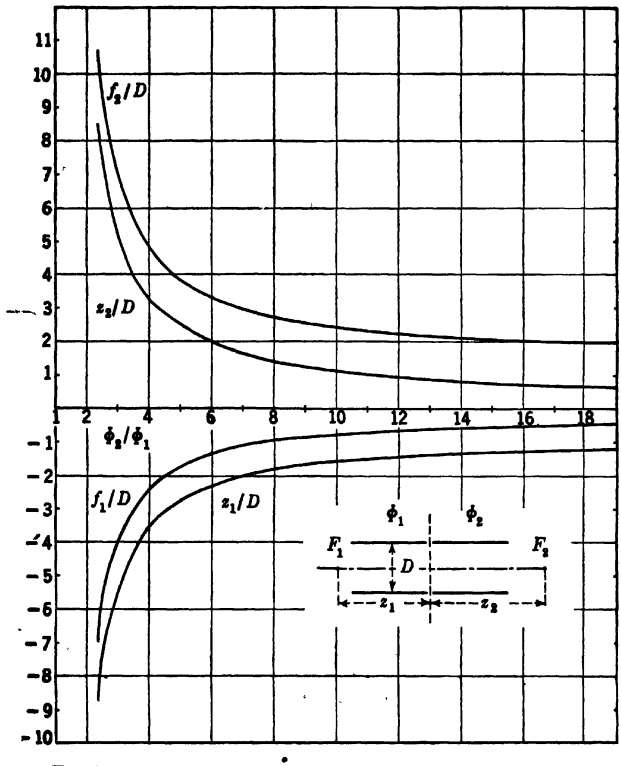


FIG. 13·21. Focal Properties of the Equidiameter Cylinder Lens.

tential distribution on the axis is given by Eq. 11·83:

$$\Phi(z) = \frac{\Phi_1 + \Phi_2}{2} + \frac{\Phi_2 - \Phi_1}{\pi} \int_0^\infty \frac{\sin \dfrac{2kz}{D}}{k \cdot J_0(ik)}\, dk \qquad [13\cdot49]$$

This is very closely approximated by[14]

$$\Phi(z) = \frac{\Phi_1 + \Phi_2}{2} + \frac{\Phi_2 - \Phi_1}{2} \tanh \frac{2.630z}{D} \qquad [13\cdot50]$$

With the aid of these formulas it is possible, by integration of the ray equation, to determine the focal lengths of such lenses as well as the distances of the focal points from the plane of symmetry. In this man-

[14] See Bertram, reference 7, and Plass, reference 8.

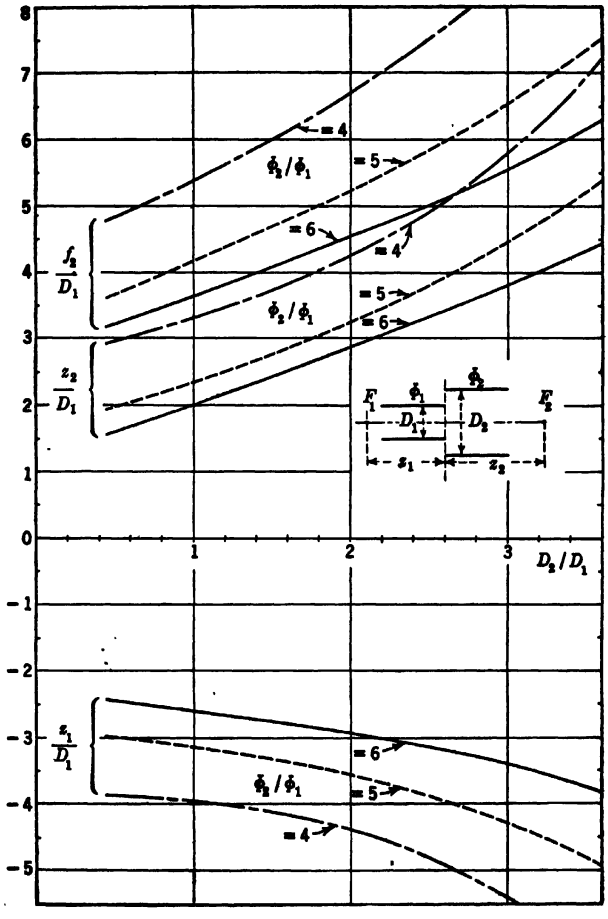


FIG. 13·22. Coaxial Cylinder Lens: Variation of Optical Parameters with Diameter Ratio. (After Epstein.)

ner the curves in Fig. 13·21 were obtained. It is seen that the crossed principal planes always lie on the low-voltage side of the immersion lens. For the two-aperture lens this follows from the form of Eqs. 13·47.

If the diameters of the two cylinders differ, purely analytical methods cease to be effective in the determination of the potential distribution. Such cases have been studied, however, by Epstein[15] with the aid of the

[15] See Epstein, reference 9.

electrolytic tank and subsequent numerical integration of the electron paths. Figure 13·22 shows the variation of the optical parameters of the lens with the diameter ratio (treating the radius of the first cylinder as the unit of length) for various voltage ratios as derived from his determinations. As is to be expected, the refractive power of the lens decreases as the diameter of the second cylinder is increased, that of the first being maintained constant.

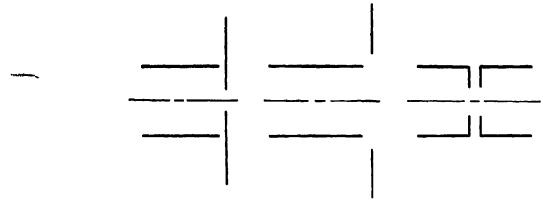


FIG. 13·23. Electrostatic Lenses Consisting of Cylinders and Apertures.

Other immersion lenses may be formed by combinations of cylinders and apertures, such as are shown in Fig. 13·23. In order to get a rough idea of the position of the cardinal points in such cases, the apertures may generally be replaced by cylinders of equal diameter. For more accurate information it is necessary to resort to the electrolytic tank and numerical integration.

13·9. Cathode Lenses. If electrons left a cathode with exactly zero velocity, cathode lenses would present no focusing problem, since all electrons from one object point would travel along the same course. Hence, if the wave properties of electron beams were disregarded, an infinitely sharp electron image would exist in all planes beyond the cathode. In fact, however, electrons leave with randomly distributed, small initial velocities. Consequently there is a preferred plane in which a true image of definite magnification is formed. As brought out in section 12·1, the location of this plane may be found by considering electron pencils leaving the cathode with infinitesimally small lateral initial velocities. In a cathode lens the image-side focal point and principal plane have their usual significance, provided that the cathode is plane. However, the cardinal points of cathode lenses possess little importance, since the object position (and, hence, the location of the sharp image) is unique.

Among cathode-lens systems, also, circular apertures or cylinders in front of a flat cathode form the simplest examples. Both result in negative lenses. In particular, the focal length of the field about a very

small aperture in a plane electrode a distance d in front of a plane cathode is, by Eq. 13·44, just $-4d$. As the aperture is increased in size, the refractive power increases and the focal length decreases. Thus

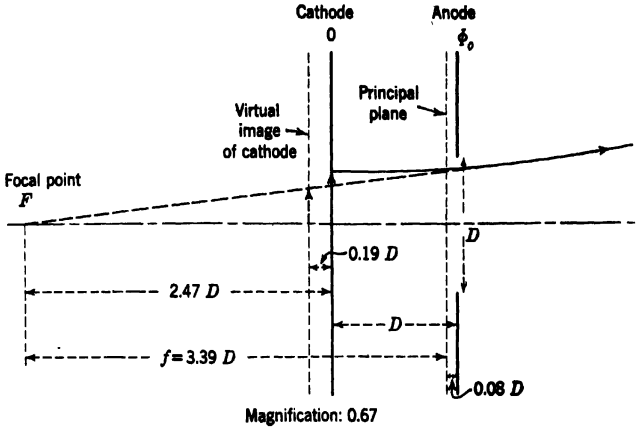


FIG. 13·24. Electron-Optical Properties of an Aperture in Front of a Plane Cathode.

(Fig. 13·24), if the aperture diameter is just equal to the distance between the two electrodes, the focal length is only $-3.39d$. The position of the virtual image of the cathode, which may be determined with the aid of the convergence equations given in section 12·1, is only $0.19d$

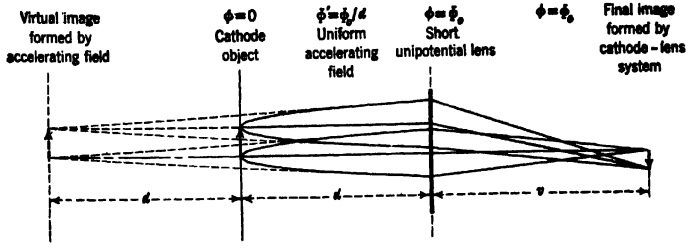


FIG. 13·25. Schematic Representation of a Cathode Lens by a Uniform Accelerating Field and a Short Unipotential Lens.

behind the cathode plane. The image magnification is 0.67. By comparison, if the aperture is very small, the lens action of the system corresponds to that of a uniform accelerating field of length d terminated by a short lens of focal length $4d$. The electrons reaching the short lens

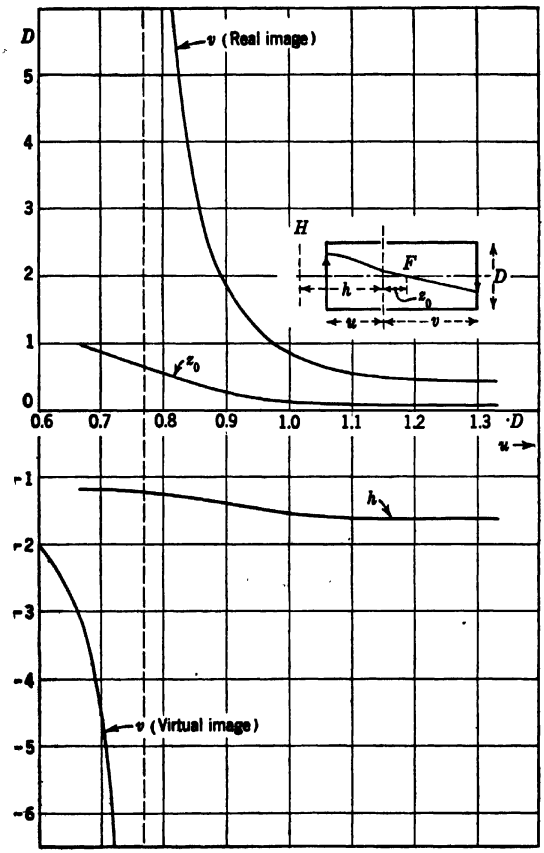


FIG. 13·26. Focal Properties of a Two-Cylinder Cathode Lens (Electrostatic Image Tube).

will thus appear to come from an object at a distance $2d$ from the lens, the object size being the same as that of the original cathode object (Fig. 13·25). By applying the lens equation, the image distance v from the lens will be given by

$$\frac{1}{v} = -\frac{1}{2d} - \frac{1}{4d} \qquad v = -\frac{4d}{3}$$

and the magnification by

$$m = -\frac{v}{2d} = \frac{2}{3}$$

Thus in this limiting case the magnification is also 0.67 and the position of the image is $0.33d$ behind the cathode in place of $0.19d$.

As has already been mentioned, a positive cylinder placed in front of a flat cathode disk also constitutes a negative lens. However, if a cylinder of equal diameter at cathode potential is interposed between the cathode disk and the positive cylinder, a real image is formed, provided that the length of the added cathode cylinder exceeds $0.76D$, D being the cylinder diameter. Figure 13·26 shows the variation of the distances of the image, the focal point, and the principal plane from the junction between the two cylinders as function of the length of the cathode cylinder, or the " object distance," u. From these three curves it may be shown that, to a good degree of approximation, the magnification is given by

$$m = \frac{v}{2u} \qquad [13·51]$$

Here v is the distance of the image from the plane separating the two cylinders. This is exactly the relation which is obtained if the cathode lens is approximated by a uniform field of length u terminated by a unipotential short lens (Fig. 13·25).

In order to be able to vary the strength of the cathode lens of the type last described, it is convenient to replace the cathode cylinder by a series of rings connected by resistors so as to simulate a uniformly resistive cylinder. One end of this ring assembly is connected to the cathode, the other to a variable voltage supply, placing it at a potential Φ_1. Figure 13·27 shows a plot of the image distance and magnification of such a system as function of the ratio Φ_1/Φ_2, Φ_2 being the potential of the anode cylinder. The object distance u is here maintained at the constant value $0.9D$. The image distance is reduced rapidly as Φ_1 is made more negative, the lens becoming increasingly convergent. The magnification, however, first decreases and then increases. This is due

to the fact that the lens not only becomes stronger, but moves progressively closer to the cathode as Φ_1 is made more negative. With increasingly positive Φ_1 the point is soon reached at which the field ceases to form a real image of the cathode.

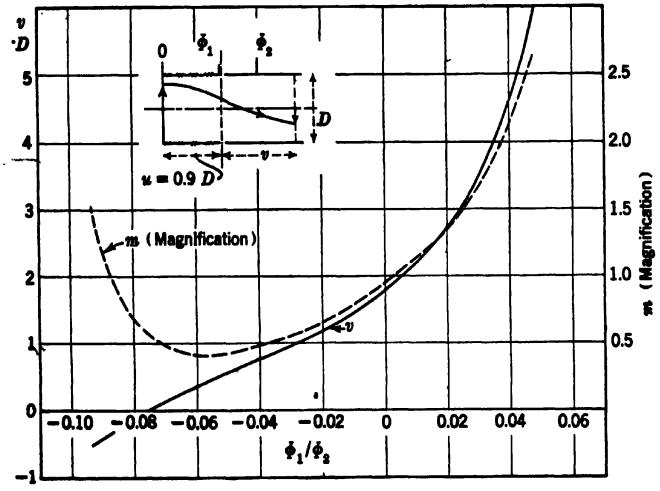


FIG. 13·27. Image Distance and Magnification for a Cathode Lens with a Resistive Cathode Cylinder.

13·10. Electron Mirrors. If, in any axially symmetric electrode configuration, the potential drops below zero at any point on the axis, electrons proceeding toward that region will reverse their direction. Thus the retarding field acts on the electron rays much as a mirror acts on light rays. Furthermore, the analysis of the action of axially symmetric fields on electron beams presented at the beginning of this chapter applies also in this case. An electron mirror consisting of an axially symmetric field distribution including a region on the axis with potential less than zero forms, to the first order of approximation, a true and faithful image of any object placed on the axis of the system and emitting, or irradiated by, electrons. It thus acts like a concave or convex (light) mirror which forms a real or virtual image of an object similarly disposed relative to it. Figure 13·28 shows a few ways of realizing electron mirrors. Generally, any electrostatic lens may readily be converted into an electron mirror by suitably altering the potentials of some of the field-forming electrodes.

As simplest example of an electron mirror, examine the uniform re-

tarding field. Such a field, it is clear, acts as a plane mirror, since a pencil of rays parallel to the field returns along its own path, all rays reversing direction in the same plane normal to the field. The plane of the equivalent mirror lies twice as far from the object plane as the plane of reversal. This is seen readily as follows (Fig. 13·29): Let the distance between object plane and plane of reversal be d, so that the retarding field is $E = \Phi/d$, Φ being the volt velocity of the electrons in the object plane. Electrons passing through a point P of the object plane with the same velocity, but at different angles θ with respect to the field direction, will describe parabolas, returning to the object plane with the same angle θ relative to the field. The time of passage from the object plane back to the object plane will be given by $2T$, where

$$d \cos^2 \theta = \frac{eE}{2m} T^2 = \frac{e\Phi}{2md} T^2 \qquad [13·52]$$

$$T = \left(\frac{2m}{e\Phi}\right)^{1/2} d \cos \theta \qquad [13·53]$$

Hence the point of intersection of the path with the object plane will be separated from P by a distance

$$y = \left(\frac{2e\Phi}{m}\right)^{1/2} \sin \theta \cdot 2T = 4d \sin \theta \cos \theta \qquad [13·54]$$

For small angles θ the virtual image is thus located a distance $4d$ from the object, just as though a plane mirror had been placed at a distance $2d$ from the object. Furthermore, this virtual image is of the same size as the object and has the same orientation as for an ordinary plane mirror. It is not perfectly sharp, however, as the image position depends, through the factor $\cos^2 \theta$, on the inclination of the imaging rays.

Even in this simplest of examples the behavior of the electron mirror is thus more complex than that of its light-optical analogue. This applies even more to convex and concave electron mirrors, as, throughout, the complex fields in front of the zero-equipotential surface play quite as important a part in the performance of the mirror as the shape of this equipotential surface itself. Thus the mirror consisting of two adjoining equidiameter cylinders at potentials of opposite sign with a plane zero-equipotential surface midway between them acts not as a plane mirror, but as a strongly convex mirror (Fig. 13·30).

It is also evident that the usual ray equations cannot be used throughout in tracing a paraxial ray through a mirror, since the slope of the ray must become infinite at the point of reversal. This difficulty can be avoided by introducing the time as parameter, both the axial and radial velocity components remaining finite throughout.

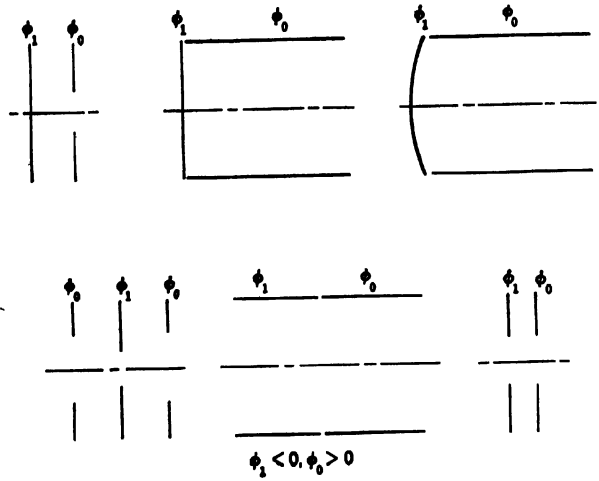


FIG. 13·28. Various Forms of Electron Mirrors.

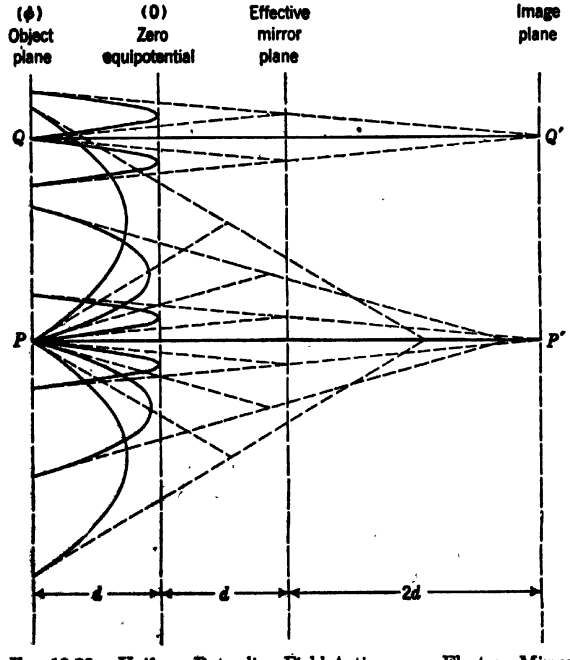


FIG. 13·29. Uniform Retarding Field Acting as an Electron Mirror.

Neglecting quantities of higher order than the first in r, the equations of motion for an electron in a meridional plane may be written

$$\dot{z} = \Phi^{\frac{1}{2}} \qquad [13\cdot55]$$

$$\ddot{r} = \frac{1}{2}\frac{\partial \phi}{\partial r} = -\frac{r}{4}\Phi'' \qquad [13\cdot56]$$

Here a dot indicates differentiation with respect to the time, measured in units of $(m/[2e])^{\frac{1}{2}}$. The relation between the time parameter and the distance along the axis can be obtained from Eq. 13·55 by simple

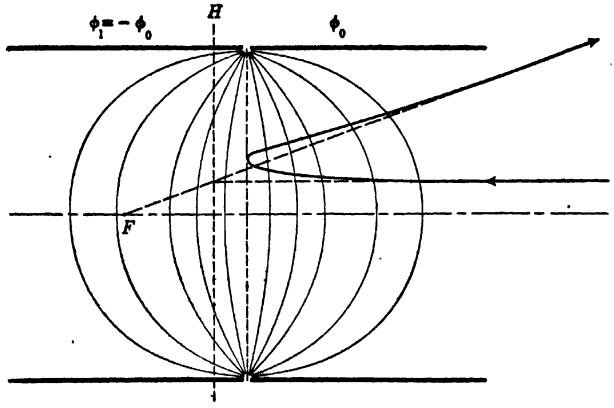


FIG. 13·30. Reflection of an Electron Ray by a Mirror Formed by Two Equidiameter Cylinders at Opposite Potentials.

quadrature. Thus, if z_u is the point of reversal, t_u the time of reversal, and z_0 is the value of the z-coordinate of the electron at the time $t = 0$,

$$t = \int_{z_0}^{z} \frac{dz}{\Phi^{\frac{1}{2}}} \qquad t < t_u \qquad [13\cdot57a]$$

$$t_u - t = \pm \int_{z}^{z_u} \frac{dz}{[\Phi_u'(z - z_u)]^{\frac{1}{2}}} = \mp 2\left[\frac{z - z_u}{\Phi_u'}\right]^{\frac{1}{2}} \qquad t \approx t_u \qquad [13\cdot57b]$$

$$t = 2t_u - t(z) \qquad t > t_u \qquad [13\cdot57c]$$

$t(z)$ in the last equation represents the value of t as function of z given by the first equation. With $t(z)$, and hence $z(t)$, known, Φ'' in Eq. 13·56 may be expressed as a function of t, so that this equation may be integrated numerically in the manner outlined in section 12·3.

It is also possible to avoid the introduction of a parameter altogether.

In this case the convergence equation (Eq. 12·14)

$$c' = c^2 - \frac{\Phi'}{2\Phi}c + \frac{\Phi''}{4\Phi} \qquad [13\cdot58]$$

is integrated to some point a small distance, z_2, from the point of reversal, $z = 0$ ($\Phi(0) = 0$ — for simplicity in notation the origin of the z-coordinate is made to fall on the plane of reversal) by the usual methods.

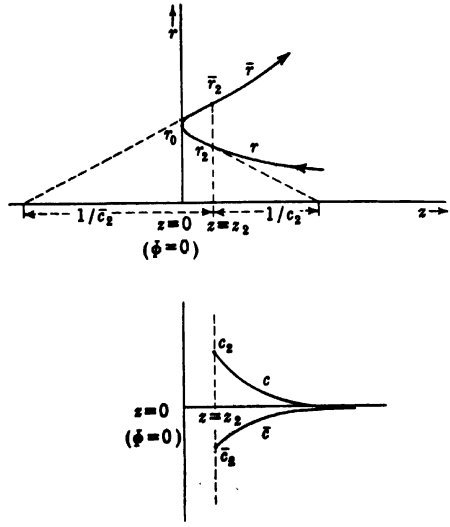


FIG. 13·31. Path Determination in an Electron Mirror.

Let $r(z_2) = r_2$, $r'(z_2) = r'_2$, $c(z_2) = c_2$, $c'(z_2) = c'_2$. If z_2 is small enough and $z = 0$ is not a saddle point of the potential, the field between $z = 0$ and $z = z_2$ may be regarded as uniform. Within it the electron describes a parabolic path with the vertex at $z = 0$. If r_o is the ordinate of the path at $z = 0$, the equation of the path in this region is

$$r - r_o = 2r'_2(z_2 z)^{1/2} \qquad [13\cdot59a]$$

$$\bar{r} - r_o = -2r'_2(z_2 z)^{1/2} \qquad [13\cdot59b]$$

Here r refers to the approach, $\bar{r}$ to the return path of the electron (Fig. 13·31).

Since, furthermore,

$$r_o = r_2 - 2r'_2 z_2 \qquad \bar{r}'_2 = -r'_2 \qquad [13\cdot60]$$

$$\bar{c}(z_2) = -\frac{\bar{r}'_2}{\bar{r}_2} = -\frac{c_2}{1 + 4c_2 z_2} \qquad [13\cdot61]$$

The derivative of the convergence becomes

$$c_2' = -\frac{c_2(1 - 2c_2z_2)}{2z_2} \qquad \bar{c}_2' = \frac{c_2(1 + 6z_2c_2)}{2z_2(1 + 4c_2z_2)^2} \qquad [13\cdot62]$$

Thus if the path has been integrated from the right (Fig. 13·31) up to z_2, r_2, the integration may be continued from the left to the right, using the initial values $\bar{c}_2$, $\bar{c}_2'$, and $\bar{r}_2 = r_2(1 + 4c_2z_2)$.

It has been found in practice that the use of the parametric equations involves, for equal accuracy of the results, less effort than the procedure just described.

Figure 13·32 shows in detail the behavior of a paraxial ray incident parallel to the axis on two simple mirrors. Both are formed by the same electrode configuration, a pair of apertures separated by a distance equal to their diameter. In the first, "convex," mirror the negative electrode has a potential numerically equal to that of the positive electrode. In the second, strongly convergent, example, the potential of the negative electrode is only one-ninth as large as that of the positive electrode. The variation of the focal length and the position of the cardinal points with the ratio of the electrode potentials for this same electrode configuration are shown in Fig. 13·33. The mirror is found to change from a diverging to a converging mirror when the potential of the negative electrode is reduced to about a fourth of that of the positive electrode. The principal plane, at the same time, recedes farther and farther behind the reversing plane.

Data regarding the variation of the refractive power of a (convergent) mirror formed by two coaxial equidiameter cylinders with the electrode-voltage ratio will be found in Fig. 17·27. This form has been studied experimentally by Nicoll.[16]

If the point of reversal is a saddle point of the potential or very close to it, as for instance in the case of an equipotential lens with its center element made just sufficiently negative to turn back all the electrons incident on it, the electron paths become quite complex and the procedure outlined in Eqs. 13·58 to 13·62 can no longer be applied. This case, important from the point of view of the control of electron currents by negative grid apertures, but not suitable for the formation of images, has been treated in detail by Recknagel.[17] It is found that the electrons execute, in the neighborhood of the saddle point, an oscillation across the axis, the number of loops increasing as the point of reversal approaches the saddle point. In the limit, the case of a lens or mirror of infinite refractive power is approached. The aberra-

[16] See reference 10.

[17] See reference 11.

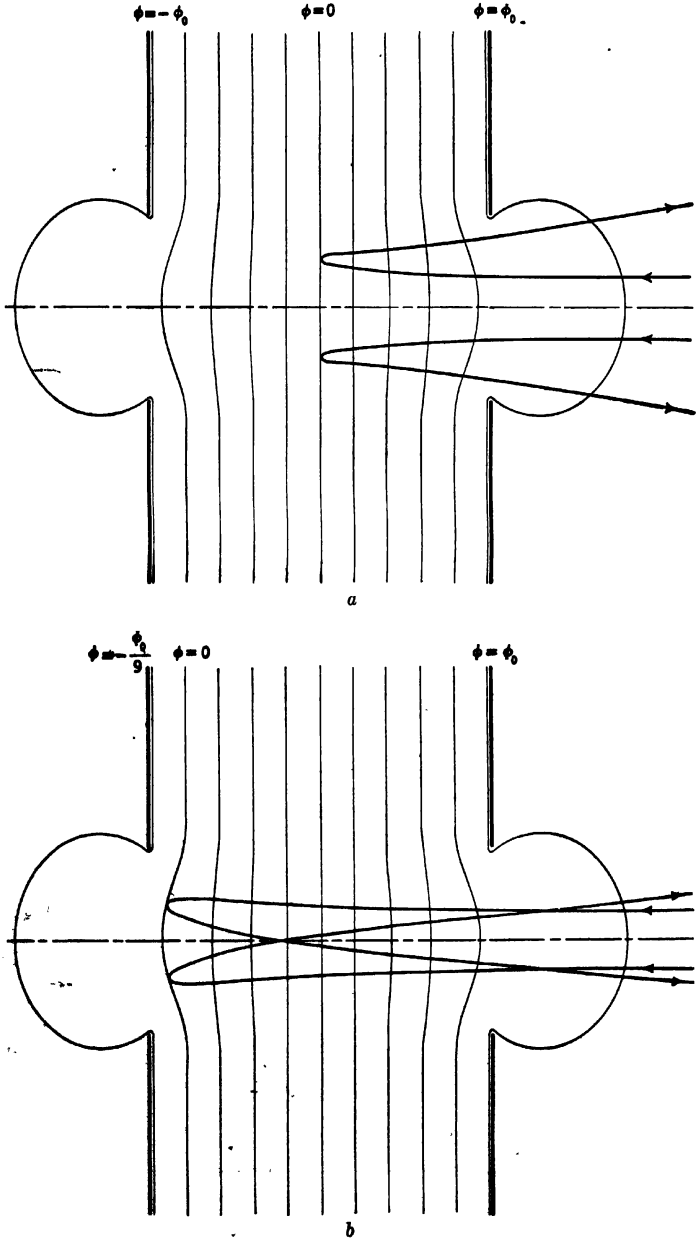


FIG. 13·32. Ray Paths in Simple Electron Mirrors.

tions of images formed by lenses or mirrors approaching this condition are so large that the images become quite unrecognizable.

In the more important case in which the field at the point of reversal ($z = z_u$) is far from zero ($\Phi_u' \neq 0$) it is possible to obtain an approximate formula for the refractive power of the mirror in much the same manner

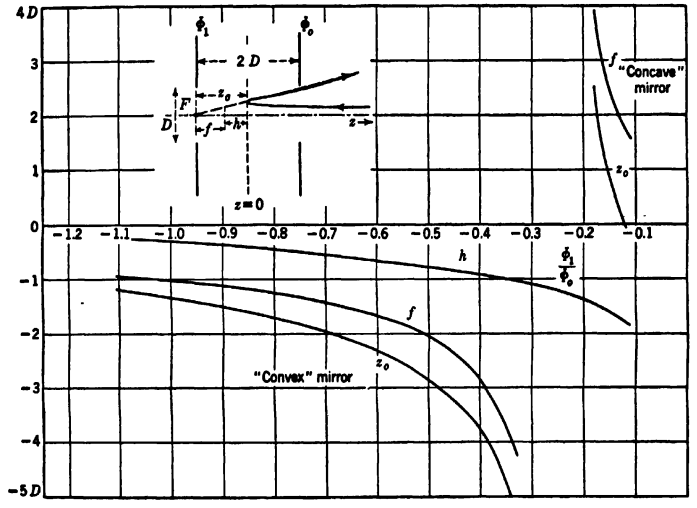


FIG. 13·33. Optical Constants of a Simple Electron Mirror as Function of the Ratio of the Electrode Potentials.

as for a thin lens (section 13·5). Let $r(z)$ be the ray approaching the point of reversal and $\bar{r}(z)$ the same ray leaving it. Then the ray equation written in the form

$$\frac{d}{dz}(r'\Phi^{1/4}) = -\frac{\Phi''}{4\Phi^{1/4}}r \qquad [13\cdot63]$$

may be integrated, for a ray incident parallel to the axis, to yield

$$(r'\Phi^{1/4})_u = -\int_{-\infty}^{z_u}\frac{\Phi''}{4\Phi^{1/4}}r\,dz \qquad [13\cdot64]$$

Although r' becomes infinite at the point of reversal and $\Phi_u = 0$, the product $r'\Phi^{1/4}$ is seen to remain finite, as the denominator in the integrand passes to zero only as $z^{1/4}$. At $z = z_u$ it reverses sign. Thus for the portion of the ray leaving the mirror,

$$\bar{r}'\Phi^{1/4} = \int_{-\infty}^{z_u}\frac{\Phi''}{4\Phi^{1/4}}r\,dz - \int_{z_u}^{z}\frac{\Phi''}{4\Phi^{1/4}}\bar{r}\,dz \qquad [13\cdot65]$$

If Φ_o is the potential at large distances from the mirror, the expression for its refractive power thus becomes

$$\frac{1}{f} = -\frac{\bar{r}'}{r_o} = \frac{1}{4\Phi_o^{1/2}} \int_{z_u}^{\infty} \frac{\Phi''}{\Phi^{1/2}} \frac{r + \bar{r}}{r_o} \, dz \qquad [13\cdot66]$$

A first approximation may be obtained by setting $\bar{r} = r = r_o$. A consideration of the actual ray paths in Figs. 13·30 and 13·32 shows, however, that this will be very greatly in error. A much better approximation results from setting $\bar{r} + r = 2r_u$, where r_u, the height of the ray at the point of reversal, is given, according to Eq. 13·64, by

$$r_u \cong r_o \left(1 - \int_{\infty}^{z_u} \frac{dz}{\Phi^{1/2}} \int_{\infty}^{z} \frac{\Phi''}{4\Phi^{1/2}} dz\right) \qquad [13\cdot67]$$

This approximation is equivalent to treating the electron as though it traversed the field of the mirror with a constant radial velocity, equal to that at the point of reversal. Substituting Eq. 13·67 in Eq. 13·66, there follows for the refractive power of the mirror the approximate relation

$$\frac{1}{f} = \frac{1}{2\Phi_o^{1/2}} \int_{z_u}^{\infty} \frac{\Phi''}{\Phi^{1/2}} dz - \frac{1}{8\Phi_o^{1/2}} \int_{z_u}^{\infty} \frac{\Phi''}{\Phi^{1/2}} dz \cdot \int_{z_u}^{\infty} \frac{dz}{\Phi^{1/2}} \int_{z}^{\infty} \frac{\Phi''}{\Phi^{1/2}} dz \qquad [13\cdot68]$$

To indicate the degree of accuracy that may be expected with the employment of Eq 13·68, it may be mentioned that the focal length of the mirror shown in Fig. 13·30 was found to be 0.34 diameter when calculated by this equation. Equation 13·66 with $\bar{r}$ and r set equal to r_o yielded 0.55 diameter. The correct value, obtained by numerical integration of the differential equations, is 0.29 diameter. The error incurred will, of course, vary greatly with the character of the mirror considered. In general, the relatively small amount of extra labor involved in integrating the parametric equations (13·55 and 13·56) will be amply compensated by the greater reliability of the results obtained.

REFERENCES

1. E. P. Adams, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, Smithsonian Institution, Washington 1922.
2. M. Knoll and E. Ruska, "Contribution to geometric electron optics," *Ann. Physik*, Series 5, Vol. 12, pp. 607–661, February 1932.
3. E. Brüche, "Acceleration of thread rays," *Z. Physik*, Vol. 78, pp. 26–42, September 1932.
4. B. v. Borries and E. Ruska, "Short space-charge field of auxiliary discharge as concentration lens for cathode rays," *Z. Physik*, Vol. 76, pp. 649–654, 1932.
5. M. Knoll and H. Weichardt, "Focal length and image quality of circular aperture lens with central screen electrode," *Z. Physik*, Vol. 110, pp. 233–236, August 1938.

6. C. J. Davisson and C. J. Calbick, " Electron lenses," *Phys. Rev.*, Vol. 38, p. 585, August 1931; Vol. 42, p. 580, November 1932.
7. S. Bertram, " Determination of the axial potential distribution in axially symmetric fields," *Proc. Inst. Radio Engrs.*, Vol. 28, pp. 418–420, September 1940.
8. G. N. Plass, " Electrostatic electron lenses with a minimum of spherical aberration," *J. Applied Phys.*, Vol. 13, pp. 49–55, January 1942; Vol. 13, p. 524, August 1942.
9. D. W. Epstein, " Electron-optical system of two cylinders as applied to cathode-ray tubes," *Proc. Inst. Radio Engrs.*, Vol. 24, pp. 1095–1139, August 1936.
10. F. H. Nicoll, " The focusing properties of the electrostatic field between two cylinders," *Proc. Phys. Soc.*, London, Vol. 50, pp. 888–898, October 1938.
11. A. Recknagel, " On the theory of the electron mirror," *Z. Physik*, Vol. 104, pp. 381–394, February 1937.

CHAPTER 14

MAGNETIC FIELDS

14·1. Determination of Fields in the Absence of Iron.[1] Magnetic fields surround any conductor carrying electric current as well as any magnetized material. The laws governing them closely resemble those governing electric fields.

The two basic equations[2] of the static magnetic field are

$$\text{div } \mathbf{B} = \nabla\cdot(\mu\mathbf{h}) = 0 \qquad [14\cdot1]$$

and

$$\text{curl } \mathbf{h} = \nabla \times \mathbf{h} = \frac{4\pi\mathbf{j}}{c} \qquad [14\cdot2]$$

Here $\mathbf{B}(=\mu\mathbf{h})$ denotes the magnetic induction, $\mathbf{h}$ the magnetic field, and μ the permeability of the medium at the point considered. In vacuum, and to a good approximation in all nonferromagnetic materials, $\mu = 1$ and $\mathbf{B} = \mathbf{h}$. $\mathbf{j}$ denotes the current density at the point of the field in question. Equation 14·2 may be converted, by integration over a surface and application of Stokes's theorem, into Ampère's law:

$$\oint \mathbf{h}\cdot d\mathbf{s} = \frac{4\pi i}{c} \qquad [14\cdot3]$$

stating that the line integral of the magnetic field about any closed

[1] Iron here refers to any material which can be magnetized to an appreciable degree.

[2] The differential operator ∇ has, in rectangular coordinates, the components $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$ and may generally be treated as an ordinary vector. Thus, if $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ denote unit vectors along the three coordinate axes, its product with a scalar, its scalar product with a vector, and its vector product with a vector, as well as the Laplacian operation ∇^2, are given by:

$$\text{grad } \psi = \nabla\psi = \mathbf{i}\frac{\partial\psi}{\partial x} + \mathbf{j}\frac{\partial\psi}{\partial y} + \mathbf{k}\frac{\partial\psi}{\partial z}$$

$$\text{div } \mathbf{h} = \nabla\cdot\mathbf{h} = \frac{\partial h_x}{\partial x} + \frac{\partial h_y}{\partial y} + \frac{\partial h_z}{\partial z}$$

$$\text{curl } \mathbf{h} = \nabla \times \mathbf{h} = \mathbf{i}\left(\frac{\partial h_z}{\partial y} - \frac{\partial h_y}{\partial z}\right) + \mathbf{j}\left(\frac{\partial h_x}{\partial z} - \frac{\partial h_z}{\partial x}\right) + \mathbf{k}\left(\frac{\partial h_y}{\partial x} - \frac{\partial h_x}{\partial y}\right)$$

$$\nabla^2\psi = (\nabla\cdot\nabla)\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}$$

path is equal to $4\pi/c$ times the current passing through any surface bounded by the path.

In any region free of electric currents and ferromagnetic materials, Eqs. 14·1 to 14·3 become

$$\nabla \cdot \mathbf{h} = 0 \qquad [14\cdot4]$$

$$\nabla \times \mathbf{h} = 0 \qquad [14\cdot5]$$

and

$$\oint \mathbf{h} \cdot d\mathbf{s} = 0 \qquad [14\cdot6]$$

These are identical in form with the equations of the electric field in the absence of space charge (Eqs. 11·1 and 11·2). Hence there exists a magnetic scalar potential ψ,

$$\mathbf{h} = -\nabla\psi \qquad [14\cdot7]$$

obeying the Laplace differential equation

$$\nabla^2\psi = 0 \qquad [14\cdot8]$$

In spite of this exact analogy between magnetic and electric fields in free space, the methods of determining them differ, in general, materially. Whereas, in the case of electric fields, the potentials of conductors bounding the region of interest are normally known and serve to determine the field distribution, the scalar magnetic potentials of surfaces surrounding a magnetic field are known only in exceptional cases. More commonly the field has to be found from given current distributions or the known magnetizations of permanently magnetized materials, which correspond to permanently polarized dielectrics. In this section magnetic fields produced by currents in the absence of iron will be considered exclusively.

Under the circumstances mentioned above it is not surprising that the scalar magnetic potential plays only a secondary role. It proves convenient, however, to introduce another quantity, the vector potential $\mathbf{A}$, with the differential properties

$$\nabla \times \mathbf{A} = \mathbf{h} \qquad [14\cdot9]$$

$$\nabla \cdot \mathbf{A} = 0 \qquad [14\cdot10]$$

to simplify the calculation of magnetic fields produced by currents. Introducing Eq. 14·9 in Eq. 14·2 and making use of Eq. 14·10 lead to

$$-\nabla \times (\nabla \times \mathbf{A}) = \nabla^2\mathbf{A} = -\frac{4\pi\mathbf{j}}{c} \qquad [14\cdot11]$$

Equation 14·11 may be regarded as a Poisson equation applying individually to each of the three vector components of **A**. It may thus be solved in perfect analogy with the usual scalar Poisson equation by an application of Green's theorem,[3] leading to

$$\mathbf{A} = \int \frac{\mathbf{j}}{cr}\, d\tau \qquad [14\cdot12]$$

Here r signifies the separation of the volume element of space $d\tau$ with the current density **j** from the point of reference. The integration is to be carried out over all space, or at least over that part of space which contains the currents responsible for the magnetic field in question.

Equation 14·12, together with Eq. 14·9, makes it possible, in principle, to calculate any stationary magnetic field in the absence of iron, provided that the distribution of the field-producing currents is known. However, many simple cases can be treated more directly by applying procedures antedating the introduction of the vector potential.

The first of these, known as the law of Biot-Savart, can be obtained simply by taking the " curl " of Eq. 14·12:

$$\mathbf{h} = \nabla \times \mathbf{A} = \int \nabla \times \left(\frac{\mathbf{j}}{rc}\right) d\tau \qquad [14\cdot13]$$

As the differential operation ∇ refers to the coordinates of the point of reference only; Eq. 14·13 may be written

$$\mathbf{h} = \int \left[\nabla\left(\frac{1}{r}\right)\right] \times \frac{\mathbf{j}}{c}\, d\tau = -\int \frac{\mathbf{r} \times \mathbf{j}}{cr^3}\, d\tau \qquad [14\cdot14]$$

If the total current i is confined to a thin wire,

$$\mathbf{j}\, d\tau = i\, \mathbf{ds} \qquad [14\cdot15]$$

where **ds** is an element of the wire. Thus in this case

$$\mathbf{h} = -i \int \frac{\mathbf{r} \times \mathbf{ds}}{cr^3} \qquad [14\cdot16]$$

The law of Biot-Savart is usually stated in the differential form:

$$\mathbf{dh} = i \frac{\mathbf{ds} \times \mathbf{r}}{cr^3} = \frac{i\, ds}{cr^2} \sin\theta \qquad [14\cdot17]$$

[3] See Jeans, reference 1, pp. 154–184.

Here θ is the angle between **ds** and **r**, and **dh** is normal to the plane containing these two vectors, as shown in Fig. 14·1.

The second method is Ampère's. As is shown by Eq. 14·3, no unique values of scalar magnetic potential can be ascribed to points in a region containing a conductor carrying current. Every integration about a loop linking current i adds an increment of $4\pi i/c$ to the potential. This indefiniteness can be removed by passing a separation surface through the current-carrying conductor, making it impossible for any closed loop to link electric current. Figure 14·2 shows possible choices of separation surfaces in the case of an infinitely long current-carrying conductor and a simple loop. By Eq. 14·3 the potential on the two sides of the separation surface must differ by $4\pi i/c$. Ampère postulated that the magnetic field due to the current i is identical with that produced by the separation surfaces considered as a magnetic double layer with a magnetic moment i/c per unit area directed normal to the surface.

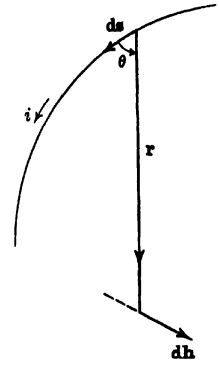


FIG. 14·1. The Law of Biot-Savart.

The validity of this theorem can be deduced simply from the law of Biot-Savart. Imagine the shell corresponding to an arbitrary current loop to be subdivided into elementary squares (Fig. 14·3). If the current i is assumed to flow in the counterclockwise direction about every one of these squares, the net effect is that of a current i around the outer boundary of the shell, currents in opposite directions canceling each other

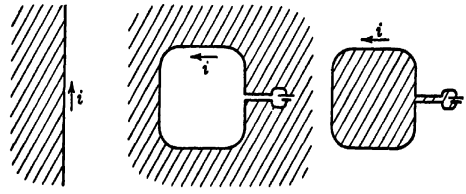


FIG. 14·2. Equivalent Magnetic Sheaths for Infinitely Long Straight Conductor and Simple Current Loop.

on the inner boundaries. Hence, also, the magnetic field due to the current around the outer boundary must be identical with the vector sum of the fields due to the elementary currents. If, for convenience, the x- and y-axes are chosen parallel to the edges of a particular element at

$P_o(x_o,y_o,z_o)$, the field due to the current about this element at the reference point $P(x,y,z)$ will be, by Eq. 14·17,

$$\begin{aligned}\mathbf{dh} &= \frac{i}{c}\left\{\frac{\mathbf{dy}\times(\mathbf{r}+\mathbf{dx})}{|\mathbf{r}+\mathbf{dx}|^3} - \frac{\mathbf{dy}\times\mathbf{r}}{|r|^3} - \frac{\mathbf{dx}\times(\mathbf{r}+\mathbf{dy})}{|\mathbf{r}+\mathbf{dy}|^3} + \frac{\mathbf{dx}\times\mathbf{r}}{|r|^3}\right\} \\ &= \frac{i}{c}\left\{-2\frac{\mathbf{dx}\times\mathbf{dy}}{r^3} - 3\mathbf{dy}\times\mathbf{r}\cdot\frac{x\,dx}{r^5} + 3\mathbf{dx}\times\mathbf{r}\cdot y\frac{dy}{r^5}\right\} \qquad [14\cdot18] \\ &= \frac{i}{c}\frac{d}{dz}\frac{\mathbf{r}}{r^3}\,dx\,dy\end{aligned}$$

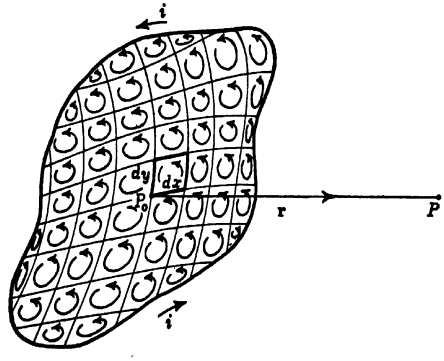


FIG. 14·3. Derivation of Ampère's Theorem from the Law of Biot-Savart.

The final expression, however, as may readily be verified, is the field due to a dipole with a magnetic moment $(i/c)\cdot dx\cdot dy = idA/c$, where dA is the area of the element considered, oriented in the z-direction — normal to the separation surface.

The concept of the equivalent magnetic sheath leads to a very simple expression for the scalar magnetic potential due to a current loop. The scalar potential of a dipole of magnetic moment $m\cdot l$ (m = pole strength, l = length) oriented in the n-direction is

$$\psi = l\frac{d}{dn}\left(\frac{m}{r}\right) = \frac{lm}{r^2}\cos\theta \qquad [14\cdot19]$$

where θ is the angle between n and r, the radius vector to the reference point P. Applying this result to an element of the magnetic shell bounded by the current i, n representing the normal to the shell, results in

$$d\psi = \frac{idA\cos\theta}{cr^2} = \frac{id\Omega}{c} \qquad [14\cdot20]$$

Here $d\Omega$ is the solid angle intercepted by dA at P. Integrating, the

scalar potential due to the current i becomes

$$\psi = \frac{i}{c}\Omega \qquad [14\cdot21]$$

That is, the scalar potential at P due to a current i is equal to the product of i/c and the solid angle Ω subtended by the loop at P.

The example of an infinite thin straight wire carrying current i (Fig. 14·4) will illustrate the application of the three procedures. Let the z-axis of the coordinate system coincide with the wire and let the reference point P be located at $(r,\theta,0)$. Then Eq. 14·12 yields for the vector potential component in the z-direction

$$A_z = \int_{-\infty}^{\infty} \frac{i\,dz_o}{c(r^2 + z_o^2)^{1/2}} \qquad [14\cdot22]$$

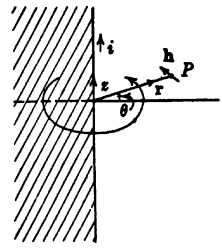


FIG. 14·4. Magnetic Field Due to Infinite Straight Wire Carrying Current.

The remaining components vanish. Accordingly the magnetic field is in an azimuthal direction and equal to

$$h_\theta = -\frac{\partial A_z}{\partial r} = \frac{ir}{c}\int_{-\infty}^{\infty} \frac{dz_o}{(r^2 + z_o^2)^{3/2}} = \frac{2i}{cr} \qquad [14\cdot23]$$

The law of Biot-Savart, Eq. 14·16, leads to the same result more directly:

$$h_\theta = \frac{i}{c}\int_{-\infty}^{\infty} \frac{r\,dz_o}{(r^2 + z_o^2)^{3/2}} = \frac{2i}{cr} \qquad [14\cdot24]$$

To apply the method of the equivalent magnetic sheath, imagine the sheath to lie in the infinite half plane $\theta = \pi$. This will cut a lune of vertex angle $-\theta$ or of solid angle -2θ out of an infinitely large sphere about P, so that, by Eq. 14·21,

$$\psi = -\frac{2i\theta}{c} \qquad [14\cdot25]$$

and, again,

$$h_\theta = -\frac{1}{r}\frac{\partial \psi}{\partial \theta} = \frac{2i}{cr} \qquad [14\cdot26]$$

The relative convenience of the three methods varies greatly from problem to problem, so that no general conclusion should be drawn from this example regarding their preferability.

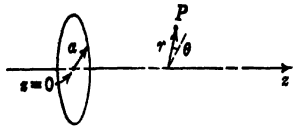


FIG. 14·5. Magnetic Field of Circular Current Loop.

The original method of the vector potential will be applied to one other problem because of its basic importance for iron-free magnetic lenses: the simple circular current loop. If the center of the loop is placed at the origin and its axis taken as the z-axis (Fig. 14·5), the vector potential will, for reasons of symmetry, be in an azimuthal direction and be given by (from Eq. 14·12):

$$A_\theta = \frac{i}{c}\int_0^{2\pi} \frac{a\cos\theta_o\, d\theta_o}{(z^2 + r^2 + a^2 - 2ar\cos\theta_o)^{1/2}} \qquad [14\cdot27]$$

Changing the variable of integration to $\theta_o' = (\pi - \theta_o)/2$ and rearranging terms result in

$$A_\theta = \frac{2i}{cr}\frac{(z^2 + r^2 + a^2)}{(z^2 + [r+a]^2)^{1/2}}\int_0^{\pi/2} \frac{d\theta_o'}{(1 - k^2\sin^2\theta_o')^{1/2}} - \frac{2i}{cr}(z^2 + [r+a]^2)^{1/2}\int_0^{\pi/2}(1 - k^2\sin^2\theta_o')^{1/2}\, d\theta_o' \qquad [14\cdot28]$$

$$= \frac{2i}{cr}(z^2 + [r+a]^2)^{1/2}\left[F\left(\frac{\pi}{2},k\right) - E\left(\frac{\pi}{2},k\right)\right] - \frac{4ia}{c(z^2 + [r+a]^2)^{1/2}}F\left(\frac{\pi}{2},k\right)$$

with

$$k^2 = \frac{4ar}{z^2 + (r+a)^2} \qquad [14\cdot29]$$

F and E are the elliptic integrals of the first and second order:

$$F\left(\frac{\pi}{2},k\right) = \frac{\pi}{2}\left\{1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1\cdot3}{2\cdot4}\right)^2 k^4 + \left(\frac{1\cdot3\cdot5}{2\cdot4\cdot6}\right)^2 k^6 + \cdots\right\} \qquad [14\cdot30]$$

$$E\left(\frac{\pi}{2},k\right) = \frac{\pi}{2}\left\{1 - \left(\frac{1}{2}\right)^2 k^2 - \left(\frac{1\cdot3}{2\cdot4}\right)^2 \frac{k^4}{3} - \left(\frac{1\cdot3\cdot5}{2\cdot4\cdot6}\right)^2 \frac{k^6}{5} - \cdots\right\} \qquad [14\cdot31]$$

Hence

$$A_\theta = \frac{2\pi ia}{c(z^2 + [r+a]^2)^{1/2}}\left\{\frac{1}{2}\left(\frac{1}{2}\right)^2 k^2 + \frac{2}{3}\left(\frac{1\cdot3}{2\cdot4}\right)^2 k^4 + \frac{3}{4}\left(\frac{1\cdot3\cdot5}{2\cdot4\cdot6}\right)^2 k^6 \cdots\right\} \qquad [14\cdot32]$$

and the two magnetic field components become

$$h_z = \frac{1}{r}\frac{\partial}{\partial r}(rA_\theta) = \frac{2\pi ia}{c(z^2 + [r+a]^2)^{3/2}}\left[a + \frac{3}{2}\left(\frac{1}{2}\right)^2 (2a - r)k^2 + \frac{5}{3}\left(\frac{1\cdot 3}{2\cdot 4}\right)^2 (3a - 2r)k^4 + \cdots\right] \quad [14\cdot 33]$$

$$h_r = -\frac{\partial A_\theta}{\partial z} = -\frac{2\pi iaz}{c(z^2 + [r+a]^2)^{3/2}}\left[\frac{3}{2}\left(\frac{1}{2}\right)^2 k^2 + \frac{10}{3}\left(\frac{1\cdot 3}{2\cdot 4}\right)^2 k^4 + \frac{21}{4}\left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^2 k^6 + \cdots\right] \quad [14\cdot 34]$$

On the axis ($r = 0$) these three expressions simplify to

$$A_\theta = 0, \qquad h_r = 0, \qquad h_z = H(z) = \frac{2\pi ia^2}{c(z^2 + a^2)^{3/2}} \quad [14\cdot 35]$$

This example has been discussed in detail because the circular current loop is the elementary constituent of all iron-free magnetic electron

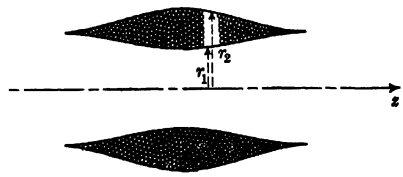


FIG. 14·6. Solenoid with Arbitrarily Distributed Windings.

lenses. Thus the axial field of an arbitrary coil (Fig. 14·6) whose radial boundaries are given by $r_1(z)$ and $r_2(z)$, and which is constructed so that there are n turns per unit cross section of the coil, is given by

$$H_1(z) = n\int_{-\infty}^{\infty} dz_o \int_{r_1(z_o)}^{r_2(z_o)} H(a, z - z_o)\, da \quad [14\cdot 36]$$

the expression for $H(a,z)$ under the integral being given by $H(z)$ in Eq. 14·35.

In the special case in which the coil consists of unevenly spaced turns of uniform radius a, with a turn distribution $n(z_o)$ centimeter^{-1}, the

field on the axis is given by[4]

$$H = \frac{2\pi i a^2}{c} \int_{-\infty}^{\infty} \frac{n(z_o)\, dz_o}{[(z - z_o)^2 + a^2]^{3/2}} \qquad [14\cdot37]$$

For an infinitely long solenoid of uniform cross section and turn distribution this yields the familiar relation

$$H = \frac{4\pi n i}{c} \qquad [14\cdot38]$$

if i is measured in electrostatic units and H in magnetic units, and

$$H = \frac{4\pi n I}{10} \qquad [14\cdot39]$$

in practical units (H in gauss, I in amperes).

It is seen that even the simplest magnetic lens field short of the uniform magnetic field is expressible, except on the axis, only in terms of the functions of higher mathematics or in series. Fortunately, it is adequate for most purposes to know the field on the axis only, together with its first few derivatives. For, since the magnetic scalar potential ψ obeys the same differential equation as the electric potential ϕ, it is possible to express its values off the axis in terms of those on the axis, in perfect analogy with Eq. 11·75:

$$\psi(r,z) = \Psi(z) - \frac{r^2}{4}\Psi''(z) + \frac{r^4}{64}\Psi^{\mathrm{IV}} \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n \Psi^{(2n)}}{(n!)^2}\left(\frac{r}{2}\right)^{(2n)} \qquad [14\cdot40]$$

The relations $h_z = -\dfrac{\partial \psi}{\partial z}$, $h_r = -\dfrac{\partial \psi}{\partial r}$ lead hence to

$$h_z(r,z) = H(z) - \frac{r^2}{4}H''(z) + \frac{r^4}{64}H^{\mathrm{IV}}(z) \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n H^{(2n)}}{(n!)^2}\left(\frac{r}{2}\right)^{2n} \qquad [14\cdot41]$$

[4] The converse problem of determining the distribution of coil windings on a cylindrical surface which will give rise to a given axial field distribution $H(z)$ has been solved by Glaser (reference 2). His formulas lead to the expression:

$$n(z_o) = \frac{c}{4\pi^2 i a} \int_{-\infty}^{\infty} \frac{e^{\frac{iuz_o}{a}}\, du}{uH_1^{(1)}(-iu)} \int_{-\infty}^{\infty} e^{-\frac{iuz}{a}} H(z)\, dz$$

Here $H_1^{(1)}$ is the Hankel function of the first kind of the first order (see, for example, Jahnke and Emde, reference 3, pp. 199 ff.). In the coefficient of the integral i signifies the current, in the exponents the imaginary quantity $(-1)^{1/2}$. u is an integration variable.

$$h_r(r,z) = -\frac{r}{2}H'(z) + \frac{r^3}{16}H'''(z) - \cdots$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^n H^{(2n-1)}}{n!(n-1)!}\left(\frac{r}{2}\right)^{2n-1} \qquad [14\cdot42]$$

It may be of interest to the reader to verify the identity of the expressions in Eqs. 14·33 and 14·34 with those obtained by substituting Eq. 14·35 in Eqs. 14·41 and 14·42.

The vector potential A can be similarly expressed if use is made of Stokes's theorem, which states that the surface integral of the normal component of the curl of any vector is equal to the line integral of the same vector about the bounding curve. If a circle of radius r about the axis is considered as surface of integration,

$$\int_0^{2\pi} A_\theta r\, d\theta = \int_0^r \int_0^{2\pi} \text{curl}_z\, A \cdot r\, dr\, d\theta = 2\pi \int_0^r h_z r\, dr \qquad [14\cdot43]$$

or

$$A_\theta = \frac{1}{r}\int_0^r h_z r\, dr \qquad [14\cdot44]$$

By substituting Eq. 14·41 in Eq. 14·44,

$$A_\theta = \frac{r}{2}H(z) - \frac{r^3}{16}H''(z) + \frac{r^5}{384}H^{IV}(z) - \cdots$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n H^{(2n)}}{n!(n+1)!}\left(\frac{r}{2}\right)^{2n+1} \qquad [14\cdot45]$$

Equation 14·44 may be written

$$A_\theta = \frac{N}{2\pi r} \qquad [14\cdot46]$$

where N is the magnetic flux through the circle of radius r passing through the reference point. This gives a simple meaning to the vector potential in axially symmetric systems.

The Eqs. 14·40 to 14·46 apply generally to all axially symmetric magnetic fields provided that the region under consideration, about and including the axis of symmetry, is free from ferromagnetic material and electric currents. It is immaterial whether electric currents or magnetic materials, or both, are the source of the field.

14·2. Magnetic Fields in the Presence of Iron. Although iron-free fields are easier to treat, they are much the less important ones in practice. Normally the use of ferromagnetic materials is valuable both for

obtaining stronger fields for a given number of ampere-turns and for shaping the field distribution in a manner suitable for the purpose at hand. Finally, the use of particular ferromagnetic materials, permanent magnets, does away with the need of electric current altogether.

To begin with, consider materials such as soft iron, mu-metal, and permalloy, which show no appreciable residual field on removal of the exciting current and thus require electric current for the production of a magnetic field. As before, the basic equations of the magnetic field are, both within and without the magnetic material,

$$\text{div } \mathbf{B} = \text{div } (\mu \mathbf{h}) = 0$$

and

$$\oint \mathbf{h} \cdot d\mathbf{s} = \frac{4\pi i}{c}$$

However, it is no longer possible to set the permeability equal to unity; for soft iron and similar substances it has values of the order of several

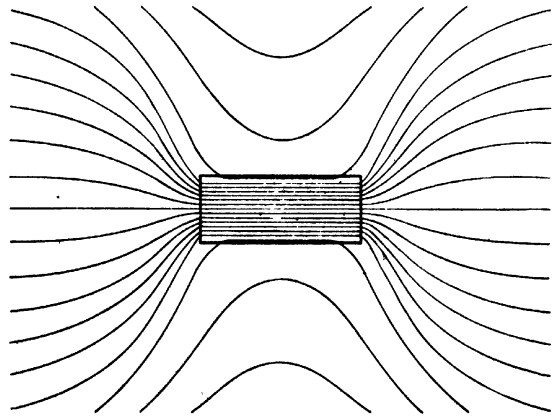


FIG. 14-7. Magnetic Material Placed in Uniform Field (Schematic).

thousands. The magnetic field $\mathbf{h}$ within the ferromagnetic material causes a degree of alignment of the molecular magnets of the material depending on the strength of the field, resulting in a magnetization per unit volume I. In a homogeneous material the net effect of this alignment is equivalent to the formation of a surface layer of magnetic poles with a surface density σ:

$$\sigma = I_n = \frac{(\mu - 1)}{4\pi} h_n \qquad [14\text{-}47]$$

where I_n and h_n are the components of the magnetization and the field normal to the surface. This polarization of the ferromagnetic material results, in general, in an enhancement of the field at the entrance of the field lines into the material and a reduction thereof within it (Fig. 14·7). As demanded by Eq. 14·1, the lines of induction

$$\mathbf{B} = \mathbf{h} + 4\pi\mathbf{I} \qquad [14\cdot48]$$

are continuous, whereas the field lines are not.

Consider, as an example (Fig. 14·8), a ferromagnetic shell, provided with a gap, about a wire carrying current i. Imagine, as on previous occasions, a separation surface to be passed through the wire, so as to make the scalar magnetic potential single-valued. The magnetic-field conditions are here identical with the electric-field conditions in a system consisting of a highly conductive shell (conductivity $\sigma = \mu$) immersed in a medium of low conductivity ($\sigma = 1$), the two halves of the shell being separated by a double-layer plane producing a difference of potential $4\pi i/c$ on the two sides. The current lines become here identical with the lines of magnetic induction, the electric-field lines with the magnetic-field lines, since the basic equations of the electrical model may be written

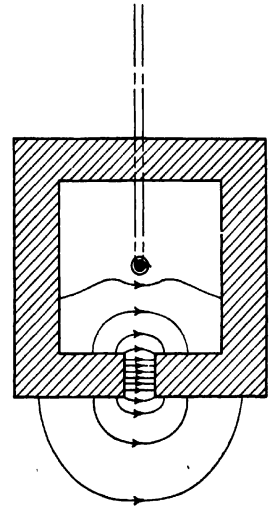


FIG. 14·8. Magnetic Circuit Formed by Magnetic Shell about Conductor.

$$\text{div } \mathbf{j} = \text{div } (\sigma\mathbf{E}) = 0 \qquad [14\cdot49]$$

$$\oint \mathbf{E}\cdot d\mathbf{s} = \frac{4\pi i}{c} \qquad [14\cdot50]$$

E is here the electric field, **j** the current density in the two media. The line integral is to be carried out from one side of the separation surface to the corresponding point on the other side. For all closed paths *within* the singly connected region it vanishes, as for the magnetic field **h**.

The analogy here given immediately suggests the electrolytic tank as a means of plotting the magnetic field. Owing to the very large value of the permeability of the iron shell compared with that of free space, the pole faces may be regarded as being at uniform potential. Furthermore, but little error results in the field map between the poles if in the model

only the region near the pole faces is immersed in a low-conductivity conducting medium, the remainder of the magnetic circuit being replaced by a conductor and voltage source (Fig. 14·9). Here the insulating wall of the electrolytic tank acts, in effect, as a reflector for the mode

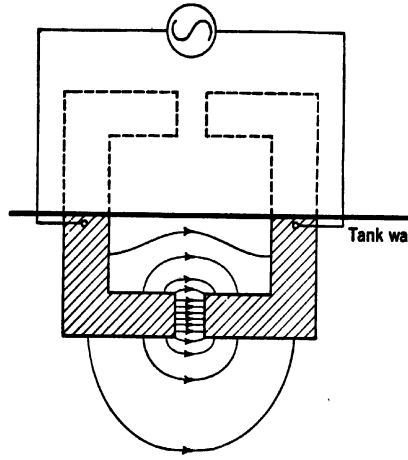


FIG. 14·9. Application of Electrolytic Tank to the Determination of the Field Distribution in the Gap between Pole Pieces.

immersed in the electrolyte. It is readily seen that for permeability ratios of the order of 1000 no attempt need be made to match the ratio of conductivities to that of the permeabilities, provided only that the difference in the conductivity of the electrodes and the electrolyte is chosen large enough.

It is obvious that the method applies equally well if the single wire in Fig. 14·8 is replaced by n parallel wires, each carrying a current i/n, so that i now stands for the sum of the currents in the individual wires. If the gap in the shell is quite large and μ is very high, not only the field distribution, but also the actual values of the field in the gap can be obtained from the field map, as the difference of magnetic scalar potential (*magnetomotive force*) is given directly by $4\pi i/c$.

If the gap is very narrow this is not true. Furthermore, an accurate field representation by the electrolytic tank may require in this case a matching of the ratio of the electrode and electrolyte conductivities to that of the permeabilities of the iron of the shell and of free space. With such a field representation at hand, the actual value of the potential difference between the equipotential surfaces in the gap can be estab-

lished by following an induction tube, that is, a tube bounded by induction or flow lines, around the magnetic circuit (Fig. 14·10). If dS is the cross section of the tube at any point, Eq. 14·3 states that

$$\oint \frac{C}{\mu dS}\, ds = \frac{4\pi i}{c} \qquad [14\cdot51a]$$

since B/μ is the density of the induction lines and $h = B/\mu$. The constant C represents the quantity BdS.

The integration is to be carried out along the axis of the induction tube. For the part of the tube within the ferromagnetic material of the shell, dS may be put equal to kS, where S is the cross section of the shell at any point and k a constant, since practically no lines of induction leave the shell except in the immediate proximity of the gap. k is given by

$$k = \frac{h\cdot dS}{\int h\cdot dS} \qquad [14\cdot51b]$$

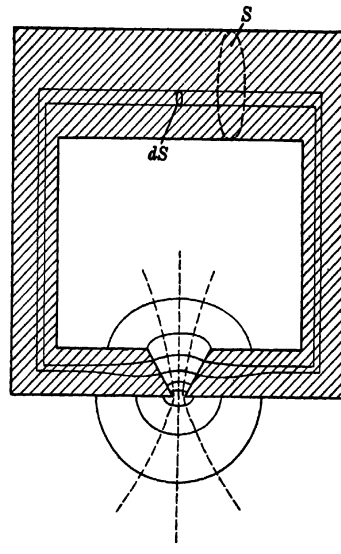


FIG. 4·10. Magnetic Field Determination about a Very Narrow Gap.

where h and dS are the magnetic field (in relative units) and the tube cross section at one particular equipotential in the gap, and the integration is carried out over that entire equipotential surface. The value of the constant C is determined for the tube in question by carrying out the integration over the circuit and substituting the result in Eq. 14·51a. The actual potential difference between any two points A and B in the tube (for example, at the faces of the gap) is given by

$$\Delta\psi = \int_A^B \frac{C}{\mu dS}\, ds \qquad [14\cdot52]$$

If this is applied to the simplest possible case of a magnetic ring of length b, permeability μ, and uniform cross-section area, with a narrow gap of width a, it is found that the magnetic difference of potential between the faces of the gap becomes

$$\Delta\psi = \frac{4\pi i a\mu}{c(b + a\mu)} \qquad [14\cdot53]$$

so that the field in the gap approaches the constant value $4\pi i\mu/(c \cdot b)$ as the gap width approaches zero.

In the discussion so far the permeability of the ferromagnetic material has been treated as a constant, independent of the exciting field h or the corresponding induction B.[5] As the curves in Fig. 14·11[6] show, this

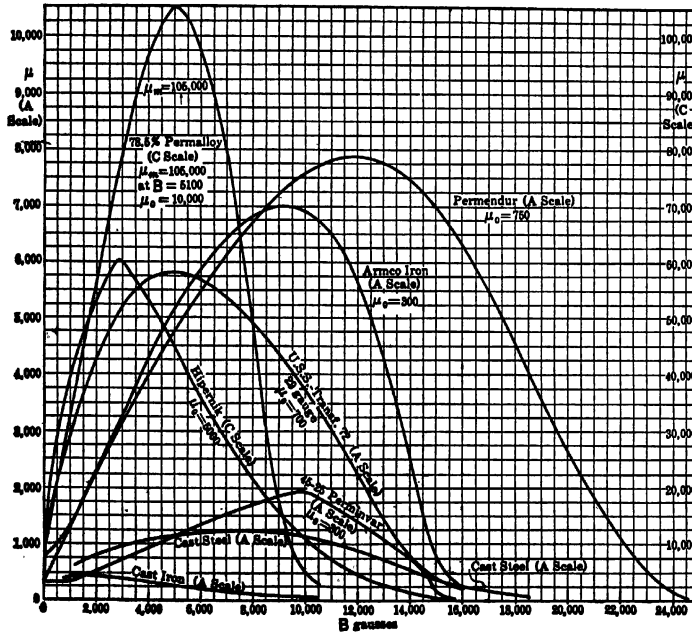


FIG. 14·11. Variation of Permeability with Magnetic Induction for Various Ferromagnetic Materials. (Reprinted from S. S. Attwood, *Electric and Magnetic Fields*, John Wiley and Sons, New York, 1941, with permission of the author.)

assumption is valid at most for limited ranges of B. In particular, the permeability drops down to a value approaching unity at very high fields; the magnetic material becomes *saturated.* Thus it is not possible to increase the field in the gap of an electromagnet indefinitely by reducing the diameter and spacing of the pole pieces or increasing the number of ampere-turns, at least not by appreciable amounts. The material of

[5] The principal conclusions, such as the applicability of the electrolytic tank, depend, fortunately, only on the requirement that, throughout, the permeability is very much greater than 1.

[6] See Attwood, reference 4, Fig. 162.

the pole faces eventually reaches a condition of maximum magnetization, and the value of its permeability drops correspondingly. Under such conditions the electrolytic tank method becomes inapplicable.

14·3. Magnetic Shielding. The fact that ferromagnetic materials placed in a magnetic field tend to concentrate the lines of magnetic induction within themselves may be utilized to screen a region of space from external magnetic fields. Such shielding is quite essential in the case of electron microscopes, as has already been indicated.[7]

Consider an infinitely long cylindrical ferromagnetic shell with inner radius r_1 and outer radius r_2 and permeability μ. Let it be placed in a uniform magnetic field with a component h_z parallel to the cylinder and one h_o normal to it. To begin with, consider the parallel component only. When the cylinder is placed in the field h_z, the symmetry conditions are such that all the lines of force must still remain parallel to the cylinder (Fig. 14·12). Taking the line integral of the magnetic field along any rectangular path with two sides parallel to the cylinder demonstrates, in view of Eq. 14·3, that the field must everywhere be identical, within and without the ferromagnetic shell. Thus the shell exerts no shielding action for the component of the magnetic field parallel to the axis of the cylinder.

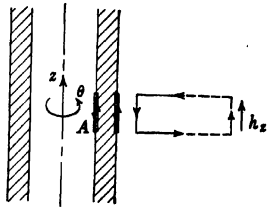


FIG. 14·12. Continuity of the Magnetic-Field Component Parallel to the Interface between Two Regions of Different Permeability.

Equation 14·3 applied to an elementary path such as A in Fig. 14·12, two sides of which are on opposite sides of any boundary surface between media of different permeability and separated by an infinitesimal distance, shows quite generally that the component of the magnetic field $\mathbf{h}$ *parallel* to the surface is the same on both sides. However, it may be concluded from Eq. 14·1 that in the case of the *normal* component it is the induction $\mathbf{B} = \mu\mathbf{h}$, and not the field, which is continuous across the surface. For, if Eq. 14·1 is integrated over an arbitrary region, Gauss's theorem states that

$$\int \text{div } \mathbf{B}\, d\tau = \int B_n\, dS = 0 \qquad [14\cdot54]$$

where S is the surface enclosing the region of integration and B_n the component of $\mathbf{B}$ normal to this surface, regarded as positive if directed outward. The application of Eq. 14·54 to an elementary flat prismatic disk (Fig. 14·13) parallel to an interface of two media, the two base

[7] See section 4·9.

surfaces being placed on opposite sides of the boundary, indicates that B_n must be the same on the two sides.

The continuity of the normal component of the induction and of the parallel component of the field will now be utilized to determine the field

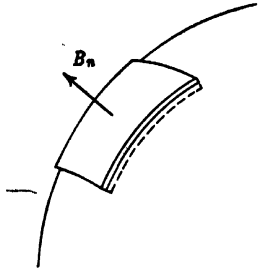


FIG. 14·13. Continuity of the Component of Magnetic Induction Normal to the Interface between Two Regions of Different Permeability.

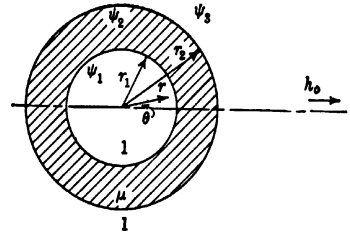


FIG. 14·14. Shielding Action of a Magnetic Shell of High Permeability.

within the cylindrical shield when placed in an originally uniform magnetic field h_o normal to the axis of the cylinder (Fig, 14·14). It is convenient to adopt polar coordinates and to denote the magnetic scalar potential functions in the shielded region, within the shield, and outside the shield by ψ_1, ψ_2, and ψ_3, respectively. Then the fact that for very large values of r the field is undisturbed may be described by

$$\lim_{|r|\to\infty} \psi_3 = h_o r \cos\theta \qquad [14\cdot55]$$

The general solution of the potential field in polar coordinates is given, according to Eq. 11·16, by the Fourier series

$$\psi = \sum_k \left(A_k r^k + \frac{B_k}{r^k}\right)(\sin k\theta + D_k \cos k\theta) \qquad [14\cdot56]$$

The symmetry conditions of the problem considered are such that all the sine terms and the cosine terms of even order must vanish. If, in addition, use is made of the fact that the potential remains finite in the central region, as well as of Eq. 14·55, the expressions for the potential in the three regions become

$$\psi_1 = b_1 r\cos\theta + b_3 r^3 \cos 3\theta + \cdots \qquad [14\cdot57]$$

$$\psi_2 = \left(c_1 r + \frac{c_1'}{r}\right)\cos\theta + \left(c_3 r^3 + \frac{c_3'}{r^3}\right)\cos 3\theta + \cdots \qquad [14\cdot58]$$

$$\psi_3 = \left(h_o r + \frac{d_1}{r}\right)\cos\theta + \frac{d_3}{r^3}\cos 3\theta + \cdots \qquad [14\cdot59]$$

The coefficients in these series are determined by the requirements of continuity of the tangential components of the field and the normal components of the induction at the interfaces:

$$\frac{\partial \psi_1}{\partial \theta} = \frac{\partial \psi_2}{\partial \theta} \qquad \frac{\partial \psi_1}{\partial r} = \mu \frac{\partial \psi_2}{\partial r} \qquad \text{at } r = r_1 \qquad [14\cdot60]$$

$$\frac{\partial \psi_2}{\partial \theta} = \frac{\partial \psi_3}{\partial \theta} \qquad \mu \frac{\partial \psi_2}{\partial r} = \frac{\partial \psi_3}{\partial r} \qquad \text{at } r = r_2 \qquad [14\cdot61]$$

As these must hold for all values of θ, they must be satisfied by the coefficients of corresponding terms in the Fourier series, leading to

$$b_1 = c_1 + \frac{c_1'}{r_1^2} \qquad b_3 = c_3 + \frac{c_3'}{r_1^6} \cdots \qquad [14\cdot62]$$

$$b_1 = \mu\left(c_1 - \frac{c_1'}{r_1^2}\right) \qquad c_3 = \mu\left(c_3 - \frac{c_3'}{r_1^6}\right) \cdots \qquad [14\cdot63]$$

$$c_1 + \frac{c_1'}{r_2^2} = h_o + \frac{d_1'}{r_2^2} \qquad c_3 + \frac{c_3'}{r_2^6} = \frac{d_3}{r_2^6} \cdots \qquad [14\cdot64]$$

$$\mu\left(c_1 - \frac{c_1'}{r_2^2}\right) = h_o - \frac{d_1'}{r_2^2} \qquad \mu\left(c_3 - \frac{c_3'}{r_2^6}\right) = -\frac{d_3}{r_2^6} \cdots \qquad [14\cdot65]$$

The solution of this simultaneous set of equations yields

$$\frac{b_1}{h_o} = \frac{4\mu}{(\mu+1)^2 - \frac{r_1^2}{r_2^2}(\mu-1)^2} \approx \frac{4r_2^2}{\mu(r_2^2 - r_1^2)} \qquad [14\cdot66]$$

Since all the coefficients of higher order than the first vanish, the field in the central, shielded region is uniform, and Eq. 14·66 gives directly the factor by which the field intensity is reduced as a result of the shielding. In regions other than the central region the field consists of a uniform field on which the field of a dipole line located along the axis of the cylinder is superposed. For the region outside the shell the dipole field is given by

$$\psi_d = \frac{d_1}{r}\cos\theta \approx -\frac{h_o r_2^2}{r}\cos\theta \qquad [14\cdot67]$$

It is, for sufficiently large values of μ and adequate thickness of the shell, practically independent of μ and corresponds to the field produced by a dipole line with a magnetic moment $-h_o r_2^2/2$ per unit length.

The same method may be applied for calculating the shielding factor for a shield consisting of several coaxial cylindrical sheaths. Thus, if a

second sheath with outer radius r_4 and inner radius r_3 is placed around the original sheath, the shielding factor becomes:

$$\frac{b_1}{h_o} \approx \frac{16r_4^2r_3^2r_2^2}{\mu^2(r_4^2 - r_3^2)(r_3^2 - r_2^2)(r_2^2 - r_1^2) + 4\mu r_3^2(r_4^2r_2^2 - r_3^2r_1^2)} \quad [14\cdot68]$$

An examination of Eqs. 14·66 and 14·68 shows that:

1. For a single shield and a given amount of material the shielding factor improves in inverse proportion to the square of the diameter of the shield.
2. With a given amount of material, a great improvement in the shielding can be obtained by subdividing the shield into several concentric cylinders separated by nonmagnetic media.

Three examples, all utilizing the same amount of material with the permeability $\mu = 10{,}000$, will illustrate this·

1. Single shell with $r_1 = 1$, $r_2 = 3.606$. Equation 14·66 yields

$$\frac{h_o}{b_1} = 2300$$

2. Single shell with $r_1 = 3.606$, $r_2 = 5$. Equation 14·66 yields

$$\frac{h_o}{b_1} = 1200$$

3. Two coaxial shells with $r_1 = 1$, $r_2 = 2$, $r_3 = 4$, $r_4 = 5$. Equation 14·68 yields

$$\frac{h_o}{b_1} = 1{,}300{,}000$$

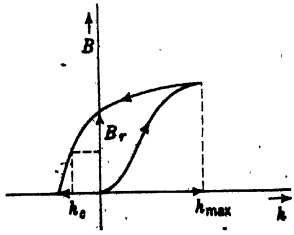


FIG. 14·15. Residual Induction B_r and Coercive Force h_c of a Permanent-Magnet Steel.

14·4. Permanent Magnets. Most ferromagnetic steels and alloys, when subjected to a strong magnetic field, do not lose their magnetism on removal of the magnetizing field, but retain a residual induction (the *remanence*) B_r. They become permanent magnets and require a considerable reversed field, the *coercive force* h_c, to render the induction within them equal to zero. Figure 14·15 shows the effect of subjecting a material of this type to an increasing field and decreasing and reversing the field after having reached some peak h_{max}.

An isolated permanent magnet (Fig. 14·16) will, in general, have a residual induction smaller than B_r. Since along any closed path, in particular along any induction line,

$$\oint \mathbf{h} \cdot d\mathbf{s} = 0 \qquad [14\cdot69]$$

the field within the magnet must be opposed in direction to that outside, the pole faces exert a demagnetizing action on the magnet. The value of the induction corresponding to the negative field h_i within the magnet may be found from the left half of the magnetization curve.

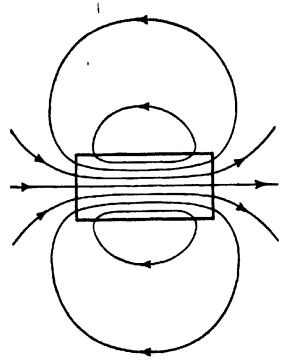


FIG. 14·16. Lines of Magnetic Induction for a Permanent Magnet (Schematic).

As has already been mentioned, permanent magnets can be used to replace coils as sources of magnetomotive force. Indeed, the fact that a coil carrying current has been shown to be equivalent in effect to a magnetic shell inserted at a convenient point in the magnetic circuit makes this immediately obvious. It is thus of interest to determine the most advantageous shaping of the magnetic material, that is, that shape which will yield the strongest field in a given gap for a given total volume of permanent-magnet material.[8]

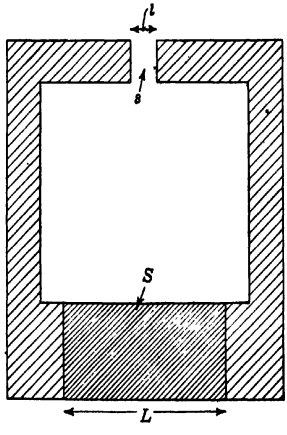


FIG. 14·17. Determination of the Most Favorable Dimensions of a Permanent Magnet.

Assume that (Fig. 14·17) the permanent magnet is of length L and cross section S and that the magnetic circuit is completed by a soft-iron armature with a gap of length l and a cross section s. Assume further, to simplify the problem, that the spreading of the lines of induction at the gap, as well as the stray field, may be neglected. Then it follows from Eqs. 14·54 and 14·69 that

$$hs = BS \qquad [14\cdot70]$$

$$hl = -h_i L \qquad [14\cdot71]$$

h and h_i being the magnetic fields in the gap and in the permanent magnet, respectively. In the second equation the contribution of the path L_a in the soft-iron armature (hL_a/μ) has been omitted because of the very large value of its permeability μ.

[8] See van Urk, reference 5.

Multiplying Eqs. 14·70 and 14·71 leads to

$$ls \cdot h^2 = -LS \cdot Bh_i \qquad [14\cdot72]$$

This means that the field strength obtainable in the gap is determined by the product of the volume ratio of the permanent magnet material and the gap and the maximum value of the product $-Bh_i$, the area of a

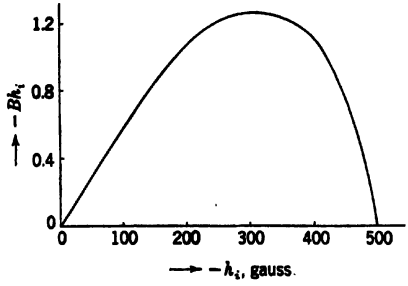


FIG. 14·18. Variation of the Product $-Bh_i$ for a Particular Permanent-Magnet Material. (van Urk, reference 5.)

rectangle inscribed under the left half of the magnetization curve as shown in Fig. 14·15. The value of $-Bh_i$ varies with the demagnetizing field h_i, that is, with the length of the permanent magnet, as shown for a particular permanent-magnet material in Fig. 14·18. The optimum value of $-h_i$ increases with the ratio of the coercive force to the remanence of the steel used. As this ratio is much greater for the modern permanent-magnet steels, such as Alnico, than for the older tungsten magnet steels, the former should be made into shorter and wider permanent magnets than the latter. If a certain value of h, the field in the gap, is prescribed, the optimum dimensions follow from Eqs. 14·70 and 14·71:

$$L = -\frac{h}{h_i} l \qquad S = \frac{h}{B} s \qquad [14\cdot73]$$

h_i and B being the values of the field and induction in the magnet, which correspond to the maximum value of $-h_iB$.

For refinements taking into account the spreading of the induction lines, the reader is referred to van Urk's original article.

14·5. Measurement of Field Distributions. As has already been shown, the calculation of magnetic fields quite generally presents considerable difficulties. In some cases, that is wherever partially saturated magnetic materials are used, it is well-nigh impossible. Also, as has been pointed out, under conditions of saturation the field determination with the aid of the electrolytic tank cannot be carried out. For this

reason special interest must be attached to methods of measuring magnetic fields directly.

Either dynamic or static methods may be employed to measure magnetic fields. The former rely on the induction of an electromotive force in a moving coil, the latter on the change of resistance in a magnetic field or on the Hall effect of certain metals such as bismuth. For a number of reasons the dynamic methods have found wider favor up to the present.

Consider a small coil of cross-section area A, number of turns n, and resistance R connected in series with a ballistic galvanometer of resistance R_g. Assume the coil to be placed in a field of strength h with its axis parallel to the field lines, so that the magnetic flux $N = nAh$ is linked by the circuit. The electromotive force induced in the circuit when the coil is moved is $-\frac{1}{c}\frac{dN}{dt}$, that is, the rate of change of the flux linked with the circuit divided by c, and the resulting current is

$$i = -\frac{1}{c}\frac{1}{R+R_g}\frac{dN}{dt} \qquad [14\cdot74]$$

If, now, the coil is suddenly pulled out of the field, a total pulse of charge

$$q = \int i dt = \frac{1}{c}\frac{N}{R+R_g} = \frac{nA}{c(R+R_g)}h \qquad [14\cdot75]$$

passes through the galvanometer, producing a deflection proportional to q and, hence, to the magnetic field h. If the damping of the galvanometer is negligible, the instantaneous deflection due to the passage of unit charge will be just $2\pi/T$ times the steady deflection for unit current, where T is the period of the galvanometer. Thus, for example, a coil of twenty turns, a cross-section area of 1 square millimeter with a resistance of 1 ohm would give, with a galvanometer of 1 ohm resistance and a period of 1 second, approximately the same deflection for a field of 1 gauss as a current of $6\cdot10^{-9}$ ampere. For the measurement of the magnetic field on the axis of an electron lens the coil is conveniently placed on an insulating rod and jerked out along this rod to produce the deflection. If the magnetomotive force is generated by electric current rather than by permanent magnets, the search coil may be left in place and the exciting current simply turned on or off or reversed. Care must be taken here to prevent error from hysteresis effects in the iron.

The sensitivity of the dynamic method can be increased, at the expense of greater complexity of apparatus, by rotating, with the aid of a synchronous motor, the coil within the field about an axis normal to the lines of force, taking off the alternating voltage by means of friction clips,

and amplifying this by means of an alternating-current vacuum-tube amplifier. If ν is the frequency of the motor, the electromotive force applied to the amplifier input becomes

$$e_i = \frac{2\pi\nu nAh}{c} \sin 2\pi\nu t = 2\pi\nu nAh\ (\sin 2\pi\nu t)\cdot 10^{-8} \text{ volt} \qquad [14\cdot76]$$

For the coil mentioned above, at 3600 revolutions per minute and $h = 1$ gauss, $e_i \approx 1$ microvolt.

Among the static methods of field measurement, that making use of the change of resistance in a magnetic field of bismuth is most valuable. Figure 14·19[9] shows the relative variation of resistance of bismuth wire and an evaporated bismuth film as a function of field strength. This property of bismuth depends greatly on the mode of preparation, shape, and even orientation in the magnetic field of the bismuth specimen, as well as on its temperature. Using flat, noninductively wound spirals of fine bismuth wire and taking extraordinary precautions to eliminate thermoelectric effects as well as other thermal effects by assembling all branches of his Wheatstone measuring bridge together on a flat strip of copper, as well as by balancing the temperature coefficients of resistance of the branches, G. S. Smith[10] has developed a reliable flux meter on the principle indicated by the diagram in Fig. 14·20. With it and a suitable bridge galvanometer a sensitivity of the

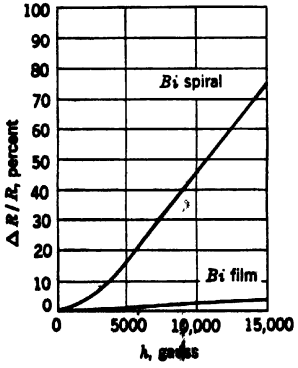


FIG. 14·19. Relative Change in Resistance of a Bismuth Spiral and a Sputtered Bismuth Film as Function of the Magnetic Field. (Bublits, reference 6. By permission of the Alien Property Custodian in the public interest under License No. A–563.)

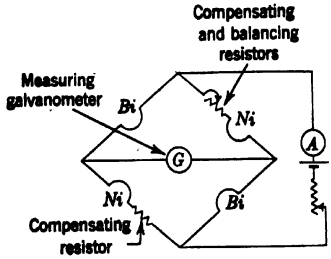


FIG. 14·20. Circuit of a Magnetic Flux Meter Employing Bismuth Spirals. (G. B. Smith, reference 7.)

[9] See Bublits, reference 6.
[10] See reference 7.

order of 20 gauss per scale division can be attained, making it well suited to the measurement of relatively high fields. The flat shape of the field probe fits it particularly for the measurement of field strengths in air gaps.

The Hall effect consists of the phenomenon that a magnetic field disturbs the current distribution in a strip of metal placed normal to it, so that a difference of potential is established in a plane normal to the original current flow. This Hall electromotive force is proportional to the current i and the magnetic field h and inversely proportional to the thickness of the strip d:

$$e_H = \frac{C \cdot h \cdot i}{d} \qquad [14\cdot77]$$

where C is a constant characteristic of the material. Figure 14·21 shows an arrangement for measuring the Hall electromotive force, and hence the magnetic field, by compensating the former with a known difference of potential.

Equation 14·77 applies only for an approximate range from 0.3 to 1000 gauss. As the magnetic field is increased beyond 1000 gauss, e_H even-

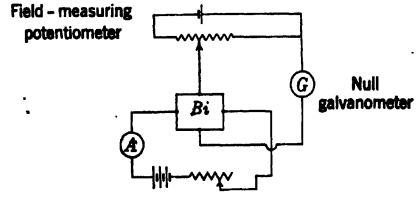


FIG. 14·21. Circuit Applying the Hall Effect of Bismuth to the Measurement of Magnetic Fields.

tually reaches a maximum and thereafter decreases, whereas for very small fields and thin films, again, the Hall constant C becomes anomalously large and the results become irregular.[11] An important drawback of the Hall effect method is its great temperature dependence.

A frequent difficulty in the application of any of the above methods of field measurement in electron optics is the small physical dimensions of the fields to be measured. Thus, in measuring the field distribution in an electron microscope objective whose pole pieces had a clear diameter of about 3.5 millimeters Dosse[12] found it necessary to construct a search coil with 100 turns of 0.02-millimeter wire whose outer diameter was 0.3 to 0.4 millimeter and whose length was 0.4 millimeter. The

[11] See Craig, reference 8.
[12] See reference 9.

displacement of the coil along the axis of the lens is measured with the aid of a microscope and an object micrometer joined rigidly to the coil. The location of the coil on the axis could be checked by the fact that, in planes normal to the axis, the field has a minimum on the axis near the center of the field distribution and a maximum on the axis in the outer portions of the lens field. For the given dimensions of the lens coil it was found that the error resulting from the spatial extent of the coil was, at the center of the field, approximately equal to the experimental error of the readings, about 0.2 per cent. Figure 14·22 shows a field distribution

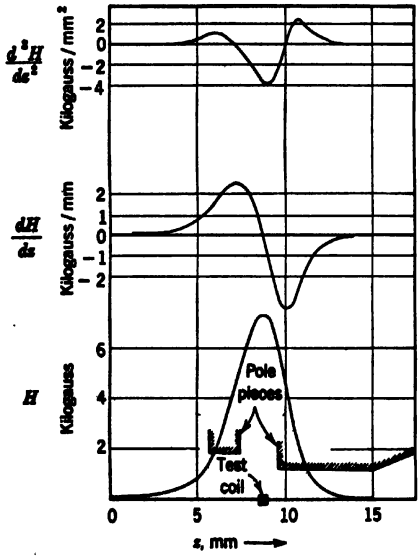


Fig. 14·22. Measured Field Distribution of a Magnetic Objective and Its Derivatives. (Dosse, reference 9.)

curve measured in this manner together with curves for the first two derivatives of the field obtained by numerical differentiation of the first curve. The number of ampere-turns in the coil was approximately 2300. Figure 14·23 shows the variation of the maximum axial field strength H_0 of the lens with the number of ampere-turns, indicating the effect of the saturation of the pole pieces above 2000 ampere-turns. The simultaneous broadening of the distribution is also shown by plotting the half-value widths of the field distribution.

Another method for measuring the axial field distribution in a magnetic electron lens utilizes the image rotation by an axially symmetric

magnetic field. The latter is given by[13]

$$\theta_i - \theta_o = \int_{z_o}^{z_i} \left(\frac{eH^2}{8m\Phi}\right)^{1/2} dz = \frac{0.148}{\Phi^{1/2}} \int_{z_o}^{z_i} H\, dz \qquad [14\cdot78]$$

so that, for fixed θ_o,

$$H(z_o) = -6.7\Phi^{1/2} \frac{d\theta_i}{dz_o} \qquad [14\cdot79]$$

Here $e\Phi$ is the kinetic energy of the beam electrons and z_o and z_i are the positions of the *object* and *image*, respectively, on the axis of the lens. The object, which may be displaced by measurable amounts along the

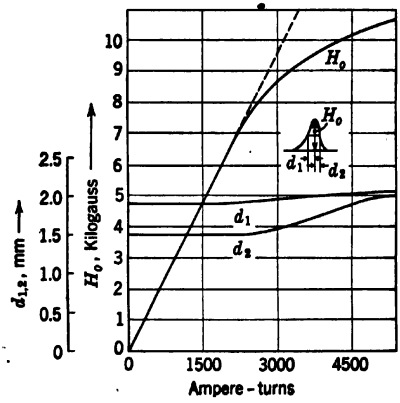


FIG. 14·23. Variation of Maximum Field Strength and Width of Distribution as Function of Number of Ampere-Turns. (Dosse, reference 9.)

axis of the lens, contains a straight edge (for example, the edge of a fine wire) which intersects the axis. It is illuminated by a beam of electrons incident parallel to the axis, which, thus, forms a shadow projection of the object on the image plane. The orientation of the straight edge (θ_i) in this image is measured for different object positions, yielding the curve $\theta_i(z_o)$, from which, by differentiation, the field distribution $H(z_o)$ may be derived. No satisfactory image pattern is obtained if z_o is near the rear focal point of the lens, since, here, the illuminating beam converges to a point. It is, therefore, necessary to interpolate values for θ_i in this region. Furthermore, the passage through the focal point is accompanied by an abrupt shift in θ_i by 180 degrees. Although this method is not suited for high-precision field measurements, it gives a rough idea of the field distribution with relatively little effort.

[13] See, e.g., Eq. 15·60.

It is possible, in principle, to circumvent the difficulties arising from the small spatial extent of the lens fields by constructing models, enlarged to scale. Care must be taken, however, if any saturating ferromagnetic material shares in the production of the field, to increase the magnetomotive force, that is, the number of ampere-turns, in the same ratio as the linear dimensions of the model.

REFERENCES

1. J. Jeans, *Mathematical Theory of Electricity and Magnetism*, Cambridge University Press, 5th Edition, 1925.
2. W. Glaser, "On the number of turns of a circular coil pertaining to a given magnetic field," *Z. Physik*, Vol. 118, pp. 264–268, October 1941.
3. E. Jahnke and F. Emde, *Tables of Functions*, B. G. Teubner, Leipzig and Berlin, 1933.
4. S. S. Attwood, *Electric and Magnetic Fields*, John Wiley and Sons, New York, 1941.
5. A. Th. van Urk, "The use of modern steels for permanent magnets," *Philips Tech. Rev.*, Vol. 5, pp. 29–35, February 1940.
6. G. Bublitz, "Measurements of constant magnetic fields with bismuth spirals and bismuth films," *Arch. tech. Messen* T 61–63 (V 391), May 1938.
7. G. S. Smith, "A new magnetic fluxmeter," *Elec. Eng.*, Vol. 56, pp. 441–445, 475–476, April 1937.
8. P. H. Craig, "The Hall effect of bismuth with low magnetic fields," *Phys. Rev.*, Vol. 27, pp. 772–778, 1926.
9. J. Dosse, "The measurement of the field of magnetic electron lenses," *Z. Physik*, Vol. 117, pp. 437–443, 1941.

CHAPTER 15

ELECTRON MOTION IN MAGNETIC FIELDS AND MAGNETIC LENSES

15·1. General Properties. Historically, the use of magnetic lenses long precedes that of electrostatic lenses. Thus, even in the Nineteenth Century a "concentrating coil" was employed by Wiechert[1] to obtain a narrow electron pencil in a cathode-ray tube. Furthermore, it was a study of the action of magnetic field coils on electron beams that led Busch[2] to formulate the general theorem that all axially symmetric magnetic and electric fields possess the properties of electron lenses.

This is not due, as might be thought at first sight, to a particularly close analogy between light optics and magnetic electron optics, but rather to the fact that magnetic fields are unimpeded by glass or metal tube walls, greatly simplifying experimentation with their action on moving electrons. Electric lenses must be built into electron beam tubes, whereas magnetic lenses may, very frequently, simply be slipped over them or wound on them. Another factor giving magnetic lenses such an important role in the beginning of electron optics is, without doubt, the easily observed ability of a uniform magnetic field to form a real image of an electron-emitting or obstructing object. A simple electric accelerating field does not possess similar properties.

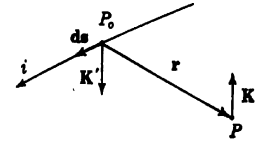


FIG. 15·1. Force Action of a Magnetic Field on a Moving Charge.

The force action of a magnetic field on a charged particle may in simple fashion be deduced from the law of Biot-Savart (Eq. 14·17) and Newton's third law, stating that to every action there is an equal and opposite reaction. According to Eq. 14·17 the force exerted on unit pole (for example, one end of a very long, thin permanent magnet) at P (Fig. 15·1) by a current i in a path element ds at P_o is

$$\mathbf{K} = i\frac{\mathbf{ds} \times \mathbf{r}}{cr^3} \qquad [15\cdot1]$$

The third law demands that an equal and opposite force must act,

[1] See reference 1.

[2] See references 2 and 3.

because of the presence of the pole, on the charges carrying the current in the element ds. If the charge in ds is dq and its mean velocity $\mathbf{v}$,

$$i\,\mathbf{ds} = \mathbf{v}\,dq \qquad [15\cdot2]$$

and the force on the charge dq is

$$\mathbf{K}' = -\mathbf{K} = -dq\,\frac{\mathbf{v}\times\mathbf{r}}{cr^3} \qquad [15\cdot3]$$

Since the magnetic field at P_o due to the unit pole at P is given by

$$\mathbf{h} = -\frac{\mathbf{r}}{r^3} \qquad [15\cdot4]$$

and since the magnetic pole can act on the charges only through the medium of its field, there results the general force law for a charge dq moving with the velocity $\mathbf{v}$ in a magnetic field $\mathbf{h}$:

$$\mathbf{K}' = \frac{dq}{c}\,\mathbf{v}\times\mathbf{h} \qquad [15\cdot5]$$

Thus for an electron of charge $-e$, moving in a magnetic field $\mathbf{h}$ and an electric field $\mathbf{E} = -\nabla\phi$, the law of motion becomes, in vector notation,

$$\dot{\mathbf{v}} = \frac{e}{m}\,\nabla\phi - \frac{e}{mc}\,\mathbf{v}\times\mathbf{h} \qquad [15\cdot6]$$

The dot over $\mathbf{v}$ indicates, in conventional manner, differentiation with respect to the time, whereas m and c represent the mass of the electron and the velocity of light, respectively. It is seen that the force exerted on the moving electron by the magnetic field is at right angles to its direction of motion, so that it changes its direction, but not its velocity. Unlike an electric field, a magnetic field does no work on a moving charge.

Optically, the effect of a magnetic field on electron beams resembles most closely the action of a doubly refracting medium on light rays. In both cases the index of refraction becomes dependent on the direction of the ray in question. As has been brought out in section 10·4, Fermat's law (Eq. 10·26) may be written, for the motion of electrons in a field with electric potential ϕ and magnetic vector potential $\mathbf{A}$ (for electron velocities small compared with that of light), in the form

$$\delta\int_M^N \left[\phi^{1/2}\,ds - \left(\frac{e}{2mc^2}\right)^{1/2}\mathbf{A}\cdot\mathbf{ds}\right] = 0 \qquad [15\cdot7]$$

the integrand representing the index of refraction n. The three Cartesian components of the vector equation 15·6 are identical with the Euler

equations

$$\frac{d}{dt}\frac{\partial F}{\partial \dot{x}} - \frac{\partial F}{\partial x} = 0 \qquad \frac{d}{dt}\frac{\partial F}{\partial \dot{y}} - \frac{\partial F}{\partial y} = 0 \qquad \frac{d}{dt}\frac{\partial F}{\partial \dot{z}} - \frac{\partial F}{\partial z} = 0 \quad [15\cdot8]$$

of Eq. 15·7 transformed into

$$\delta \int_M^N \left[\phi^{1/2} \cdot v - \left(\frac{e}{2mc^2}\right)^{1/2} \mathbf{A}\cdot\mathbf{v} \right] dt \equiv \delta \int_M^N F(x,y,z;\dot{x},\dot{y},\dot{z})\, dt = 0 \quad [15\cdot9]$$

The reader can readily verify the equivalence of Eqs. 15·8 and 15·6 if he makes the following substitutions in the former set of equations:

$$v = (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2} = \left(\frac{2e\phi}{m}\right)^{1/2} \quad [15\cdot10]$$

$$\mathbf{h} = \nabla \times \mathbf{A} \quad [15\cdot11]$$

The first of these is a consequence of the conservation of energy; the second, a part of the definition of the vector potential.

15·2. Motion in Uniform Magnetic Field. As the simplest example, consider a uniform field parallel to the z-axis, $h_z = H$. For this condition Eq. 15·6 reduces to

$$\ddot{x} = -\frac{e}{mc}\dot{y}H \quad [15\cdot12a]$$

$$\ddot{y} = \frac{e}{mc}\dot{x}H \quad [15\cdot12b]$$

$$\ddot{z} = 0$$

The last equation can be integrated directly to give

$$z = \dot{z}_o t + z_o \quad [15\cdot13]$$

where t is the time and z_o and $\dot{z}_o$ are the initial values of the z-coordinate and the z-component of velocity of the electron. The component of motion in the direction of the magnetic field is thus not affected by the latter. Correspondingly, also, the kinetic energy of the electron motion projected on the z-axis, $e\phi_z = \frac{m\dot{z}_o^2}{2}$, remains constant throughout. Since the total kinetic energy of the electron, $e\phi = e\phi_z + e\phi_{xy}$, where $e\phi_{xy} = \frac{m}{2}(\dot{x}^2 + \dot{y}^2)$ is the kinetic energy of the electron motion projected on the xy-plane, is a constant in the magnetic field, ϕ_{xy} must also be constant.

To determine the projection of the electron motion on the xy-plane,

Eqs. 15·12*a* and 15·12*b* may be differentiated with respect to the time and the values for $\ddot{x}$ and $\ddot{y}$ given by them substituted in the result:

$$\dddot{y} = -\left(\frac{eH}{mc}\right)^2 \dot{y} \qquad \dddot{x} = -\left(\frac{eH}{mc}\right)^2 \dot{x} \qquad [15\cdot14]$$

A first integration yields

$$\ddot{y} = -\left(\frac{eH}{mc}\right)^2 (y - y_c) \qquad \ddot{x} = -\left(\frac{eH}{mc}\right)^2 (x - x_c) \qquad [15\cdot15]$$

where x_c and y_c are arbitrary constants. These two equations are solved quite generally by

$$x - x_c = \rho \sin\left(\frac{eH}{mc} t + \theta_o\right) \qquad [15\cdot16]$$

$$y - y_c = \rho' \cos\left(\frac{eH}{mc} t + \theta_o'\right) \qquad [15\cdot17]$$

Here $\rho, \theta_o, \rho', \theta_o'$ are integration constants. Substitution of this solution in Eqs. 15·12*a* and 15·12*b* establishes the following relation between them:

$$\rho' = -\rho \qquad \theta_o' = \theta_o \qquad [15\cdot18]$$

Equations 15·16 to 15·18 indicate that the motion of the electron projected on the xy-plane is a circle of radius ρ, with center at the point x_c, y_c. The magnitude of the radius ρ is obtained by determining the kinetic energy corresponding to the motion described by the equations in question:

$$e\phi_{xy} = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}\left(\frac{eH}{mc}\right)^2 \rho^2 \qquad [15\cdot19]$$

whence

$$\rho = \left(\frac{2mc^2}{e}\right)^{1/2} \frac{\phi_{xy}^{1/2}}{H} = 3.372 \frac{\phi_{xy}^{1/2}}{H} \text{ cm} \qquad [15\cdot20]$$

ϕ_{xy} being measured in volts and H in gauss.

If the coordinates of the point of origin $x(0) = x_o$, $y(0) = y_o$ are introduced, Eqs. 15·16 and 15·17 pass over into

$$x = x_o + \rho\left[\sin\left(\frac{eH}{mc} t + \theta_o\right) - \sin\theta_o\right] \qquad [15\cdot21a]$$

$$y = y_o - \rho\left[\cos\left(\frac{eH}{mc} t + \theta_o\right) - \cos\theta_o\right] \qquad [15\cdot21b]$$

By differentiating it is seen that $\theta_o = \arctan(\dot{y}_o/\dot{x}_o)$, that is, that θ_o is the angle in the xy-plane which the projected electron path makes with the x-axis at the origin. Together with Eq. 15·13 for the translation in the direction of the field, Eqs. 15·21 show that the electron describes a helix whose diameter, for a given magnetic field, is proportional to its velocity component normal to the field (Eq. 15·20) and whose *pitch* d (Fig. 15·2) is determined by its velocity component parallel to the field:

$$d = \frac{2\pi mc}{eH} v_z = 2\pi c \left(\frac{2m}{e}\right)^{1/2} \frac{\phi_z^{1/2}}{H} = 21.08 \frac{\phi_z^{1/2}}{H} \text{ cm} \qquad [15·22]$$

The frequency of rotation of the electron,

$$\nu = \frac{eH}{2\pi mc} = 2.800 \cdot 10^6 H \text{ sec}^{-1} \qquad [15·23]$$

depends exclusively on the strength of the magnetic field and is independent of the velocity of the electron.

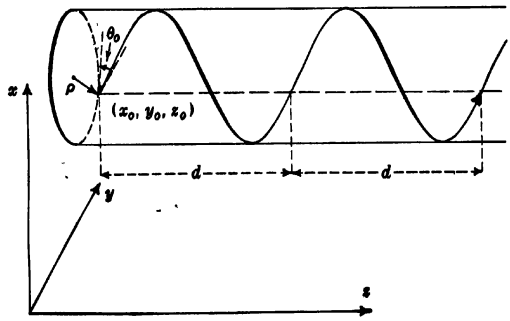


FIG. 15·2. Motion of Electron in Uniform Magnetic Field.

Consider now electrons leaving the origin O on the z-axis with the inclination σ_o. Then in Eqs. 15·21 and 15·13:

$$x_o = y_o = z_o = 0 \qquad [15·24a]$$

$$\rho = \left(\frac{2mc^2}{e}\right)^{1/2} \frac{\phi^{1/2}}{H} \sin \sigma_o \qquad [15·24b]$$

$$\dot{z}_o = \left(\frac{2e\phi}{m}\right)^{1/2} \cos \sigma_o \qquad [15·24c]$$

In cylindrical coordinates

$$x = r \cos \theta \qquad y = r \sin \theta \qquad [15·25]$$

Equations 15·21 become, in this special case,

$$r = 2\rho \sin (\theta - \theta_o) \qquad [15\cdot26a]$$

$$\theta = \frac{eH}{2mc} t + \theta_o \qquad [15\cdot26b]$$

By substituting the value of t from Eq. 15·13, utilizing Eq. 15·24c, the last equation becomes

$$\theta = \left(\frac{eH^2}{8mc^2\phi}\right)^{1/2} \frac{z}{\cos \sigma_o} + \theta_o \qquad [15\cdot27]$$

The path equations (15·26a and 15·27) may be put into a particularly convenient form by expanding in terms of σ_o:

$$r = 2\rho \sin \alpha + \Delta r \qquad [15\cdot28a]$$

$$\theta - \theta_o = \alpha + \Delta\theta \qquad [15\cdot28b]$$

where

$$\Delta r \cong 2\rho \cos \alpha \cdot \Delta\theta = z \cos \alpha \left(\frac{1}{2}\sigma_o^3 + \cdots\right) \qquad [15\cdot29a]$$

$$\Delta\theta = \alpha \left(\frac{1}{2}\sigma_o^2 + \frac{5}{24}\sigma_o^4 + \cdots\right) \qquad [15\cdot29b]$$

and

$$\alpha = \left(\frac{eH^2}{8mc^2\phi}\right)^{1/2} z \qquad [15\cdot30]$$

It is thus seen that, for angles of inclination at O which are small enough so that the terms involving the second and higher powers of σ_o may be neglected (that is, $\Delta r \cong 0$ and $\Delta\theta \cong 0$), all electrons leaving O are reunited on the axis at $O_1, O_2, \cdots O_n$ (Fig. 15·3), the separation OO_n being given by

$$\alpha_n = \left(\frac{eH^2}{8mc^2\phi}\right)^{1/2} OO_n = \pi n \qquad [15\cdot31]$$

or

$$OO_n = \pi n \left(\frac{8mc^2\phi}{eH^2}\right)^{1/2} = 21.08n \frac{\phi^{1/2}}{H} \qquad [15\cdot32]$$

Since any line parallel to the magnetic field might have been chosen as axis, the same applies to similarly spaced points Q, Q_n; S, S_n; etc., Q and S lying in the plane normal to the field passing through O. Accordingly the homogeneous magnetic field forms a sequence of uniformly

spaced real, erect images with unity magnification of an object placed normal to the field which emits, or is " illuminated " by, electrons of uniform velocity.

The existence of Δr indicates, however, that the image obtained with electrons leaving the object at finite angles to the axis is not sharp. If

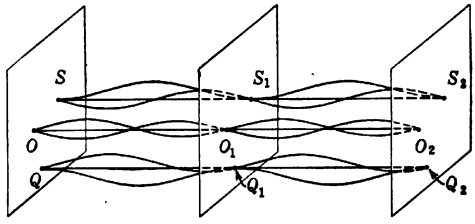


FIG. 15·3. Image Formation by a Uniform Magnetic Field.

the normal plane through O_n is considered as the image plane, electrons leaving O with an inclination σ_o will strike the image plane at a distance from the image point for *paraxial electrons*, O_n, equal to

$$\Delta r = \frac{z}{2}\sigma_o^3 + \cdots = n\pi\left(\frac{2mc^2\phi}{eH^2}\right)^{1/2}\sigma_o^3 \cdots = 10.54n\frac{\phi^{1/2}}{H}\sigma_o^3 \quad [15\cdot33]$$

This *aberration* thus increases, neglecting the fifth and higher powers of σ_o, linearly with the order n of the image and is inversely proportional, for the same order n, to the strength of the focusing field H.

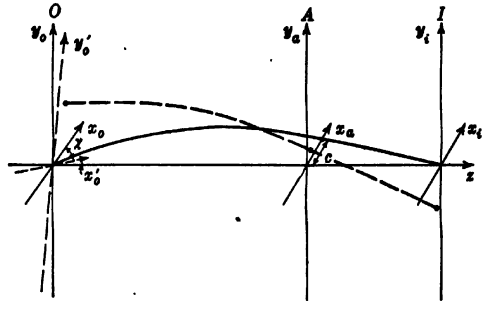


FIG. 15·4. Object, Aperture, and Image Planes for an Axially Symmetric Electric and Magnetic Field.

15·3. Motion in Axially Symmetric Magnetic Fields. As in section 13·1, consider (Fig. 15·4] a perfectly general axially symmetric configuration with a magnetic as well as an electrostatic field. Let x_o,y_o be the coordinates of an electron in the object plane, x_i,y_i those in

the image plane, and x_a,y_a its coordinates in some " aperture plane." It will be assumed that the image plane is in a field-free region and that, furthermore, the region between the image and aperture planes is field-free, so that here the electron paths are straight lines. Once again, the coordinates x_i,y_i of the intersection of an electron path with the image plane are fully determined when the intersections with the object and aperture planes, x_o,y_o and x_a,y_a, are known and x_i,y_i may be expressed as a function of these four coordinates by the power series given in Eqs. 13·1. However, the degree of symmetry of the mixed magnetic and electric field is less than that of the electric field alone. For, although any rotation about the axis leaves the field unchanged in either case, a reflection at a plane through the axis reverses the sense of the currents producing the field as well as that of the magnetic vector potential and hence the direction of the magnetic field itself. Thus the electron path and, accordingly, the coefficients in Eqs. 13·1 will in general differ for the field obtained by such a " reflection " from those for the original field. Accordingly the symmetry requirements lead to Eqs. 13·7:

$$x_i = a_1x_o + a_2y_o + a_3x_a + a_4y_a + O(3) \qquad [15\cdot34a]$$

$$y_i = -a_2x_o + a_1y_o - a_4x_a + a_3y_a + O(3) \qquad [15\cdot34b]$$

and no farther. As before, the expression $O(3)$ lumps together terms of the third and higher orders in x_o,y_o,x_a,y_a.

Consider now a ray leaving the intersection of the axis with the object plane and intersecting the aperture plane in the point $x_a = c$, $y_a = 0$. The corresponding point in the image plane is given by

$$x_i = a_3c + O(3) \qquad [15\cdot35a]$$

$$y_i = -a_4c + O(3) \qquad [15\cdot35b]$$

In a purely electric field the entire path of the electron would lie in the same meridional plane (that is, plane through the axis), since all forces exerted on electrons in such a field are confined to meridional planes. The magnetic field alone causes the electron to leave the meridional plane in which it has started. However, it can readily be shown that an electron, which anywhere within this field crosses the axis, has zero angular velocity $\dot{\theta}$ about the axis (or moves in a meridional plane) in any region where the magnetic field vanishes altogether. If the equation of motion of the electron, Eq. 15·6, is expressed in cylindrical coordinates r, θ, z, the equation expressing the variation of the angular momentum of the electron about the axis is given by

$$\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta}) = \frac{e}{mc}(\dot{r}h_z - \dot{z}h_r) \qquad [15\cdot36]$$

where h_z and h_r are the z-component and the r-component of the magnetic field, respectively. Let the coordinates and the angular velocity of the electron at the time $t = t_o$ be r_o, θ_o, z_o; $\dot{\theta}_o$ and the corresponding quantities at the time $t = t_1$ be r_1, θ_1, z_1; $\dot{\theta}_1$. Then the integration of Eq. 15·36 yields

$$r_1^2\dot{\theta}_1 - r_o^2\dot{\theta}_o = \frac{e}{mc}\int_{t_o}^{t_1} r(\dot{r}h_z - \dot{z}h_r)\,dt$$

$$= \frac{e}{mc}\left\{\int_{r_o}^{r_1} rh_z\,dr - \int_{z_o}^{z_1} rh_r\,dz\right\} \qquad [15\cdot37]$$

Substituting the expressions for h_z and h_r given by Eqs. 14·41 and 14·42 in Eq. 15·37 gives

$$r_1^2\dot{\theta}_1 - r_o^2\dot{\theta}_o = \frac{2e}{mc}\left\{\int_{z_o}^{z_1}\sum_{n=0}^{\infty}\frac{(-1)^n H^{(2n)}\left(\frac{r}{2}\right)^{2n+1}}{(n!)^2}\frac{dr}{dz}\,dz\right.$$

$$\left. - \int_{z_o}^{z_1}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}H^{(2n+1)}\left(\frac{r}{2}\right)^{2n+2}}{(n+1)!n!}\,dz\right\} \qquad [15\cdot38]$$

An integration by parts of the first term yields, in addition to the integrated sum, an integral identical except for sign with the second term of Eq. 15·38. Hence, simply:

$$r_1^2\dot{\theta}_1 - r_o^2\dot{\theta}_o = \frac{2e}{mc}\sum_{n=0}^{\infty}\left[\frac{(-1)^n H^{(2n)}\left(\frac{r}{2}\right)^{2n+2}}{(n+1)!n!}\right]_{z_o,r_o}^{z_1,r_1}$$

$$= \frac{e}{mc}\left[\frac{1}{2}Hr^2 - \frac{1}{16}H''r^4 + \frac{1}{384}H^{\mathrm{IV}}r^6 - \cdots\right]_{z_o,r_o}^{z_1,r_1} \qquad [15\cdot39]$$

If the initial point of the integration is a point on the axis, that is, if $r_o = 0$, the value of the expression on the right vanishes at the lower limit (z_o,r_o). Furthermore, if z_1,r_1 is in a region free from magnetic field, $H(z_1)$ and its derivatives vanish so that here also the expression becomes zero. Hence for the electron considered, leaving the object plane at its intersection with the axis,

$$r_1^2\dot{\theta} = 0$$

in the region between the aperture plane and the image plane, so that the electron must remain in the meridional plane through the point $x_a = c$, $y_a = 0$, that is, in the xz-plane.

Select now as the image plane the plane in which the straight line representing the electron path for $c \doteq 0$ beyond the aperture plane intersects the axis. Then $x_i = y_i = 0$ and, since $c \neq 0$, Eq. 15·35 demands that $a_3 = a_4 = 0$. Since c is assumed to be very small, the terms of third and higher order in c, represented by $O(3)$, may be neglected.

Thus Eqs. 15·34 become

$$x_i = a_1 x_o + a_2 y_o + O(3) \qquad [15\cdot40a]$$

$$y_i = -a_2 x_o + a_1 y_o + O(3) \qquad [15\cdot40b]$$

Imagine now (Fig. 15·4) the coordinate system of the object plane to be rotated through an angle χ such that

$$\tan \chi = -\frac{a_2}{a_1} \qquad [15\cdot41]$$

Then the coordinates x_o', y_o' of the object point in this coordinate system will be given by

$$x_o' = \frac{a_1 x_o + a_2 y_o}{a_1'} \qquad [15\cdot42a]$$

$$y_o' = \frac{-a_2 x_o + a_1 y_o}{a_1'} \qquad [15\cdot42b]$$

with

$$a_1' = (a_1^2 + a_2^2)^{1/2} \qquad [15\cdot43]$$

In terms of these new object plane coordinates the position of the image point is given by

$$x_i = a_1' x_o' + O(3) \qquad [15\cdot44a]$$

$$y_i = a_1' y_o' + O(3) \qquad [15\cdot44b]$$

Thus again, by neglecting terms of the third and higher orders in the coordinates of the electron paths in the object and aperture planes, a sharp, geometrically similar image of the object, either real or virtual, is produced. Its magnification is given by the factor a_1'. However, in distinction to the case of the purely electric field or a light-optical lens system, the image is rotated through any angle χ with respect to the object.

To determine this angle of rotation χ, as well as the position and magnification of the image for any particular field, it is necessary to take recourse to the equations governing the electron motion in the magnetic field. If account is taken of the fact that in axially symmetric fields

only the azimuthal component of the vector potential **A**, $A_\theta = A$ diffe from zero, Eq. 15·7 becomes

$$\delta \int F\, dz = \delta \int \left\{ \phi^{1/2} (r'^2 + r^2\theta'^2 + 1)^{1/2} - \left(\frac{e}{2mc^2}\right)^{1/2} A r\theta' \right\} dz = 0 \quad [15\cdot45]$$

Its Euler equations

$$\frac{d}{dz}\frac{\partial F}{\partial r'} - \frac{\partial F}{\partial r} = 0 \quad [15\cdot46]$$

$$\frac{d}{dz}\frac{\partial F}{\partial \theta'} - \frac{\partial F}{\partial \theta} = 0 \quad [15\cdot47]$$

accordingly take the form

$$\frac{d}{dz}\frac{r'\phi^{1/2}}{(r'^2 + r^2\theta'^2 + 1)^{1/2}} - \frac{(r'^2 + r^2\theta'^2 + 1)^{1/2}}{2\phi^{1/2}}\frac{\partial \phi}{\partial r} - \frac{\phi^{1/2} r\theta'^2}{(r'^2 + r^2\theta'^2 + 1)^{1/2}} + \left(\frac{e}{2mc^2}\right)^{1/2} \theta' \frac{\partial}{\partial r}(rA) = 0 \quad [15\cdot48]$$

$$\frac{d}{dz}\left\{ \frac{r^2\theta'\phi^{1/2}}{(r'^2 + r^2\theta'^2 + 1)^{1/2}} - \left(\frac{e}{2mc^2}\right)^{1/2} rA \right\} = 0 \quad [15\cdot49]$$

The second of these equations can be integrated immediately:

$$\frac{r^2\theta'\phi^{1/2}}{(r'^2 + r^2\theta'^2 + 1)^{1/2}} - \left(\frac{e}{2mc^2}\right)^{1/2} rA = C \quad [15\cdot50]$$

C being a constant of integration. It is seen to vanish if for any point of the path of the electron $r = 0$, that is, if the path crosses the axis, or if, in field-free space ($A = 0$), it travels in a meridional plane ($\theta' = 0$). In a space free of magnetic field it represents, furthermore, but for a constant factor $(2em)^{1/2}$, the angular momentum of the electron about the axis, since $\phi^{1/2}$ is proportional to the velocity of the electron and its coefficient is the product of r and the direction cosine for the tangential component of velocity.

To put the other equation, Eq. 15·48, into convenient form, introduce the function D:

$$D = \frac{1}{\phi^{1/2}}\left[\frac{C}{r} + \left(\frac{e}{2mc^2}\right)^{1/2} A\right] \quad [15\cdot51]$$

and solve Eq. 15·50 for $r\theta'$:

$$r\theta' = D\left[\frac{r'^2 + 1}{1 - D^2}\right]^{1/2} \qquad (r'^2 + r^2\theta'^2 + 1)^{1/2} = \left[\frac{r'^2 + 1}{1 - D^2}\right]^{1/2} \quad [15\cdot52]$$

These relations may now be used to eliminate θ' from Eq. 15·48. In addition, the term containing the vector potential A may be expressed in terms of D with the aid of the relation, following directly from Eq. 15·51:

$$\left(\frac{e}{2mc^2}\right)^{\frac{1}{2}} \frac{\partial}{\partial r}(rA) = \frac{\partial}{\partial r}(r\phi^{\frac{1}{2}}D) \qquad [15\cdot53]$$

Substituting Eqs. 15·52 and 15·53 in Eq. 15·48, carrying out the indicated differentiations, and collecting terms result finally in the differential equation

$$r'' = \frac{1 + r'^2}{2\phi(1 - D^2)}\left\{\frac{\partial[\phi(1 - D^2)]}{\partial r} - r'\frac{\partial[\phi(1 - D^2)]}{\partial z}\right\}. \qquad [15\cdot54]$$

This differential equation is identical in form with Eq. 12·9 for the path of an electron in a meridional plane of an axially symmetric electric field. Thus, provided that the initial direction of motion (determining C from Eq. 15·50) and the distribution of the magnetic field in terms of its vector potential as well as that of a superposed electric field are given, all the methods of graphical and numerical path determination given in Chapter 12 may be applied directly to plotting the variation of the separation of the electron from the axis as function of z. It is necessary only to replace the potential function ϕ by the more complex function

$$\phi(1 - D^2) = \phi - \left[\frac{C}{r} + \left(\frac{e}{2mc^2}\right)^{\frac{1}{2}} A\right]^2 \qquad [15\cdot55]$$

The meridional plane in which the electron is located at any time and in which, hence, r is measured is not the same throughout even for an electron which crosses the axis somewhere along its path. It rotates about the axis, its azimuth being given at any time (according to Eq. 15·52) by

$$\theta = \theta_o + \int_{z_o}^{z} \frac{D}{r}\left(\frac{r'^2 + 1}{1 - D^2}\right)^{\frac{1}{2}} dz \qquad [15\cdot56]$$

The equations given up to this point are exact, applying to any electrons traveling in axially symmetric fields. They can be greatly simplified if attention is confined to rays proceeding close to the axis and making small angles therewith, that is, *paraxial rays*. Substituting the expansions in r (from Eqs. 11·75 and 14·45)

$$\phi = \Phi - \frac{r^2}{4}\Phi'' + \frac{r^4}{64}\Phi^{IV} - \cdots \qquad [15\cdot57a]$$

$$A = \frac{r}{2}H - \frac{r^3}{16}H'' \cdots \qquad [15\cdot57b]$$

in Eqs. 15·54 and 15·56, and neglecting terms containing powers of r and r' higher than the first, gives the equations

$$r'' = -r'\frac{\Phi'}{2\Phi} - r\left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi} - \frac{C^2}{\Phi r^4}\right) \qquad [15\cdot58]$$

$$\theta = \theta_o + \int_{z_o}^{z}\left[\frac{C}{r^2\Phi^{1/2}} + \left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2}\right]dz \qquad [15\cdot59]$$

Here C is not a general constant of the differential equation, but one determined by the initial conditions of any one particular path. Since C, being in substance the angular momentum of the electron about the axis in field-free space, is proportional to the initial value of r^2, the equation has linear character in the sense that even with $C \neq 0$ its solutions are similar for equal initial values of θ' and r'/r. For $C = 0$, that is, for rays passing through the axis or incident on the magnetic field in a meridional plane, the angle of rotation of the electron about the axis and its path in the plane rotating with the electron are determined by the two equations

$$\theta = \theta_o + \int_{z_o}^{z}\left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2}dz \qquad [15\cdot60]$$

$$r'' = -r'\frac{\Phi'}{2\Phi} - r\left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi}\right) \qquad [15\cdot61]$$

Equation 15·61, like the analogous Eq. 12·10 for a purely electric field, may be simplified somewhat by the introduction of the dependent variable $R = r\Phi^{1/4}$:

$$R'' = -TR \qquad T = \frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 + \frac{eH^2}{8mc^2\Phi} \qquad [15\cdot62]$$

This has the same form as Eq. 12·11 for the electric field. T, in either case, is a uniformly positive function.

The identity in form of Eqs. 15·61 and 15·62 with the corresponding Eqs. 12·10 and 12·11 for the purely electric field makes it possible to apply immediately the theorems deduced in section 13·3 to mixed electric and magnetic fields. Thus

1. An arbitrary axially symmetric mixed electric and magnetic field constitutes an electron lens with object- and image-side focal points F_o, F_i, principal planes H_o, H_i, and nodal points N_o, N_i (Fig. 15·5). In the presence of a magnetic field a ray aimed at the object nodal point N_o will appear to emerge from the image nodal point N_i with its inclination to the axis unaltered. However, the meridional plane of

emergence will be rotated with respect to the meridional plane of incidence by an angle χ.

2. The focal lengths, $f_o = F_oH_o$ and $f_i = H_iF_i$, are related by $f_o/f_i = (\Phi_o/\Phi_i)^{1/2}$, where Φ_o and Φ_i are the electric potentials in the object and image spaces, respectively.

3. The nodal points are located by the relations $N_iF_i = f_o$, $F_oN_o = f_i$.

4. Both the nodal points and the principal planes are crossed (for fields forming not more than one real image); that is, H_o and N_o lie closer to F_i than H_i and N_i, respectively.

5. Negative electron lenses (free of space charge and conductors within the lens field) bounded by uniform-potential regions in object and image space do not exist.

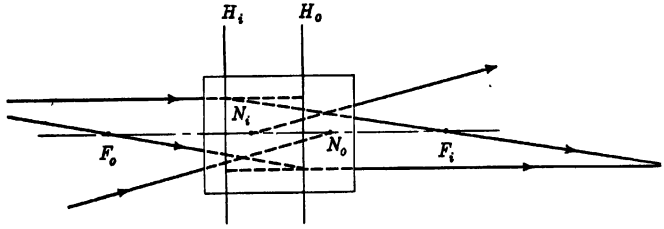


FIG. 15·5. The Cardinal Points of an Arbitrary Electron Lens. In the presence of a magnet field the ray paths in image space (right of figure) are situated in a meridional plane rotated through an angle χ with respect to the meridional plane containing the ray paths in object space (left of figure).

The analytical determination of electron paths in magnetic lens fields is effective in only a few cases. The simplest of these, the uniform field, has already been considered in section 15·1. Among the more complex fields which have been studied in this manner, the " bell-shaped " field

$$H = \frac{H_o}{1 + \left(\frac{z}{a}\right)^2} \qquad [15·63]$$

where $2a$ is the half-value width of the field distribution, has been investigated quite exhaustively by Glaser.[3] It is of special interest because of its resemblance to the magnetic fields commonly employed as electron microscope objectives. With this field the paraxial ray equation (15·61) may be written in the dimension-free form

$$y'' = -\frac{k^2 y}{(1 + x^2)^2} \qquad [15·64]$$

[3] See reference 4.

with

$$y = \frac{r}{a} \qquad x = \frac{z}{a} \qquad k^2 = \frac{eH_o^2 a^2}{8mc^2\Phi} \qquad [15\cdot65]$$

The introduction of the new independent variable β,

$$x = \operatorname{ctn} \beta,$$

brings Eq. 15·64 into the form

$$\frac{d^2y}{d\beta^2} + 2 \operatorname{ctn} \beta \frac{dy}{d\beta} + k^2 y = 0 \qquad [15\cdot66]$$

This may be solved by the substitution $y = A(\beta)\cdot v(\beta)$:

$$A\frac{d^2v}{d\beta^2} + \left(2\frac{dA}{d\beta} + 2A \operatorname{ctn} \beta\right)\frac{dv}{d\beta} + \left(\frac{d^2A}{d\beta^2} + 2\frac{dA}{d\beta}\operatorname{ctn}\beta + k^2A\right)v = 0 \qquad [15\cdot67]$$

if A is chosen so that the coefficient of $dv/d\beta$ vanishes:

$$A = \exp\left(-\int \operatorname{ctn} \beta d\beta\right) = \frac{1}{\sin \beta} \qquad [15\cdot68]$$

This leaves as equation for v:

$$\frac{d^2v}{d\beta^2} + (k^2 + 1)v = 0 \qquad [15\cdot69]$$

Hence the general solution of Eq. 15·66 becomes

$$y = A\cdot v = \frac{1}{\sin \beta}\{C_1 \sin [(k^2+1)^{1/2}\beta] + C_2 \cos [(k^2+1)^{1/2}\beta]\} \qquad [15\cdot70]$$

Consider, in particular, a ray incident from the right (positive z) parallel to the axis. For it y must remain finite for $x \doteq \infty$ ($\beta \doteq 0$) so that

$$y = \frac{1}{\sin \beta} C_1 \sin [(k^2+1)^{1/2}\beta] \qquad [15\cdot71]$$

or, in terms of the original coordinates (Eqs. 15·64 and 15·65),

$$r = C_1\left(1 + \left[\frac{z}{a}\right]^2\right)^{1/2} \sin\left\{(k^2+1)^{1/2} \arctan\frac{a}{z}\right\} \qquad [15\cdot72]$$

$$0 < \arctan\frac{a}{z} < \pi$$

The shape of these paths projected on the same meridional plane is shown, for several values of k^2, in Fig. 15·6. The field distribution has been drawn in to indicate the behavior of the electrons in the field. It is seen that for large k^2 (strong fields) the ray intersects the axis several

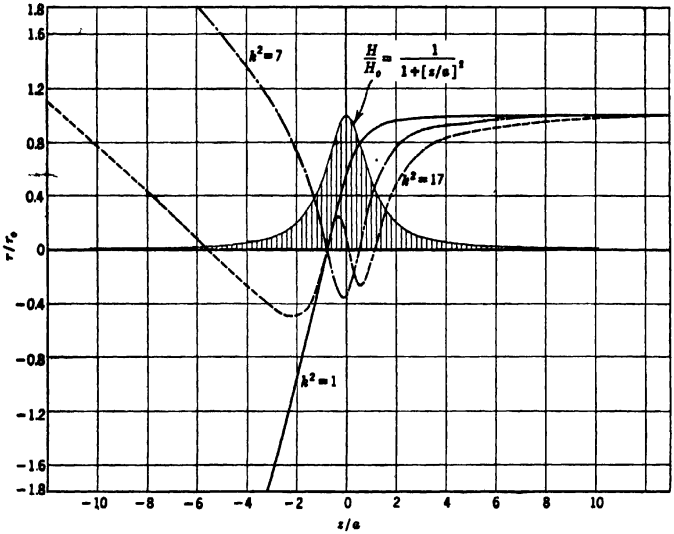


FIG. 15·6. Paraxial Electron Paths in "Bell-Shaped" Field for Various Values of $k^2 = eH_0^2a^2/(8mc^2\Phi)$.

times. As in the uniform magnetic field, the points of intersection lie closest together and the amplitude of the oscillations about the axis is least in the regions of greatest field strength. The points of intersection are given by

$$(k^2+1)^{\frac{1}{2}} \arctan \frac{a}{z_n} = n\pi \qquad z_n = a \operatorname{ctn} \frac{n\pi}{(k^2+1)^{\frac{1}{2}}} \qquad [15\cdot73]$$

$$n = 1, 2, 3, \cdots \qquad n_{\max} < (k^2+1)^{\frac{1}{2}}$$

Thus the lens has one focal point for $k^2 < 3$, two focal points for $k^2 < 8$, three focal points for $k^2 < 15$, etc. The corresponding refractive powers (for example, those of the lens used as an electron microscope objective) are given by

$$\frac{1}{f_n} = \frac{r'(z_n)}{r(\infty)} = \frac{(-1)^{n-1}}{a} \sin \frac{n\pi}{(k^2+1)^{\frac{1}{2}}} \qquad [15\cdot74]$$

The focal length can thus in no case be numerically smaller than a, half the half-value width of the distribution. Figure 15·7 shows the refractive powers and positions of the focal points as function of k^2.

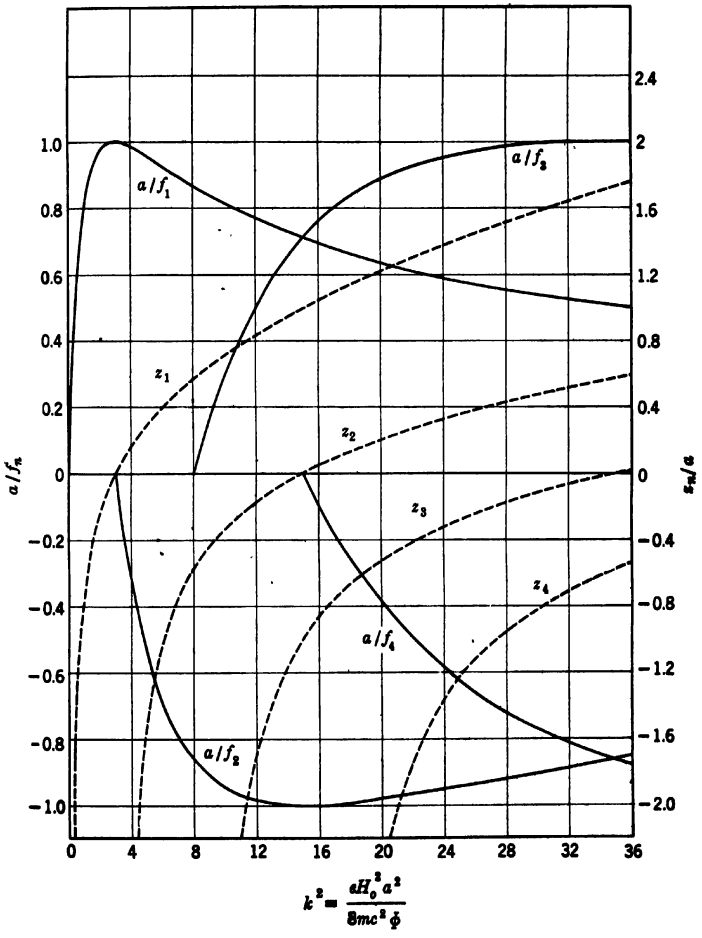


FIG. 15·7. Refractive Powers and Focal Points for Bell-Shaped Field $H/H_0 = 1/[1 + (z/a)^2]$ Employed as Objective (Object Real).

If the complete lens field is utilized for imaging, that is, if the object is virtual as in the case of the electron microscope projector,[4] the focal

[4] See section 4·5.

length becomes unique and is given by

$$\frac{1}{f} = \frac{r'(-\infty)}{r(\infty)} = -\frac{1}{a(k^2+1)^{1/2}} \sin\left[(k^2+1)^{1/2}\pi\right] \qquad [15\cdot75]$$

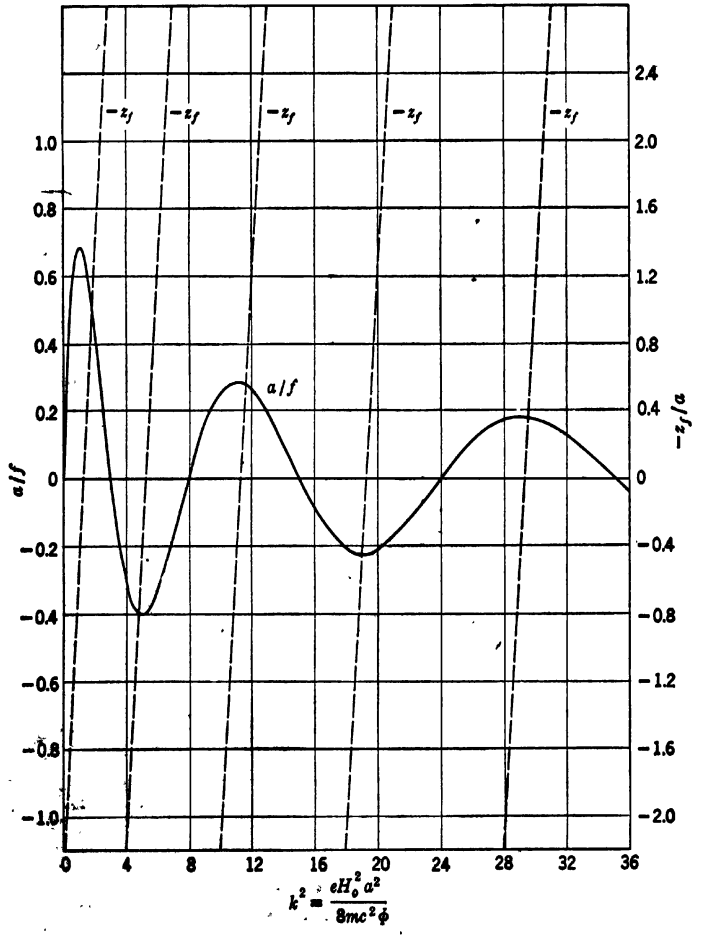


FIG. 15·8. Refractive Powers and Focal Points for Bell-Shaped Field Employed as Projector (Object Virtual).

Again, the focal length cannot be smaller than a. As k^2 is increased, the refractive power oscillates between positive and negative values with decreasing amplitude. The (virtual) object position or focal point

is given by

$$z_f = \left(z - \frac{r}{r'}\right)_{z \doteq -\infty} = -a(k^2+1)^{\frac{1}{2}} \operatorname{ctn}\left[(k^2+1)^{\frac{1}{2}}\pi\right] \quad [15\cdot76]$$

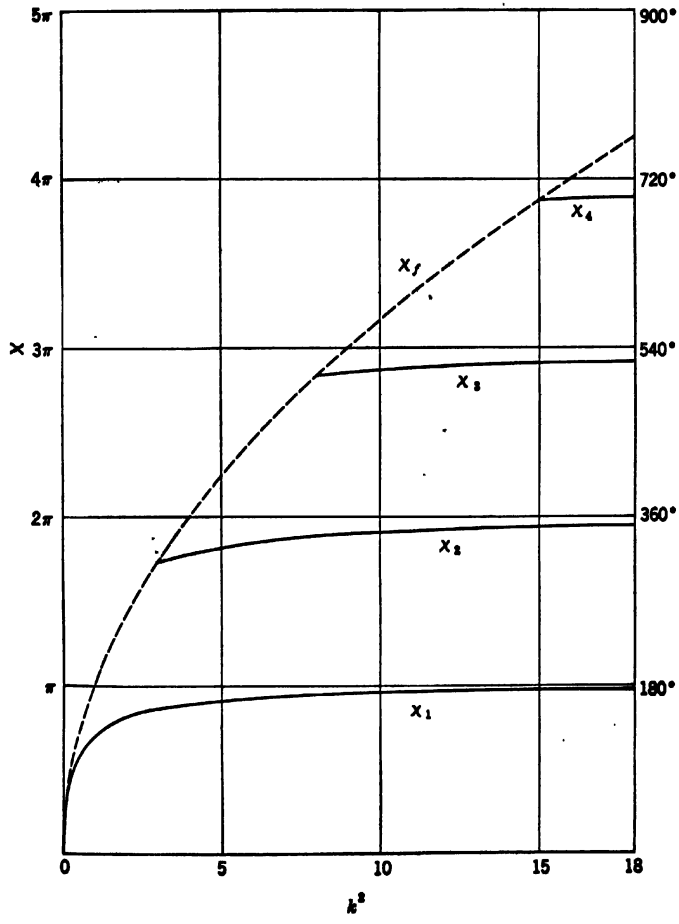


FIG. 15·9. Image Rotation Produced by the Bell-Shaped Field.

The variation of both the refractive power and the position of this focal point with k^2 is shown in Fig. 15·8. The same scale has been chosen for Figs. 15·8 and 15·7 to facilitate comparison. It is seen that the refractive powers of the lens used as projector (Fig. 15·8) are quite

generally smaller than those of the lens used as objective. The refractive powers of the projector oscillate back and forth with decreasing amplitude as the maximum field strength H_o is increased. The bell-shaped field differs from the magnetic fields normally employed as electron lenses primarily in the more gradual dropping off of the field strength at large distances from the center of the lens. This accentuates the difference in the refractive powers of the lens employed as objective and as projector.

The information derived above determines the proper object position and the magnification of the image formed by the bell-shaped field under various imaging conditions. Another quantity of interest is the image rotation, given by Eq. 15·60. Substituting Eq. 15·63 yields

$$\chi = \theta_i - \theta_o = \left(\frac{eH_o^2a^2}{8mc^2\Phi}\right)^{1/2}\left(\arctan\frac{z_i}{a} - \arctan\frac{z_o}{a}\right) \qquad [15\cdot77]$$

Substituting $z_i = \infty$ and $z_o = z_n$ (Eq. 15·73) leads to

$$\chi_n = \theta_i - \theta_o = \frac{n\pi k}{(k^2+1)^{1/2}} \qquad [15\cdot78]$$

which approaches $n\pi$ for large k^2. If the lens is used as projector, however,

$$\chi_f = \theta_i - \theta_o = \pi k \qquad [15\cdot79]$$

Both variations are shown graphically in Fig. 15·9.

The properties of a number of other magnetic electron lenses have already been discussed in sections 4·4 and 4·5. They may serve as additional examples of the action of axially symmetric fields on electron beams.

15·4. Approximate Methods of Path Determination in Axially Symmetric Magnetic Fields; the Thin Lens; Magnetic Lenses without Image Rotation. The normal procedures for determining paths in magnetic fields as well as the focal properties of magnetic electron lenses are perfectly analogous to those for finding the corresponding quantities for electrostatic lenses. The graphical methods of section 12·5, that is, those based on Snell's law and the *circle method*, can be applied directly to rays originally parallel to the axis or inclined to it in a meridional plane if, to begin with, an *equipotential map*

$$\phi(1 - D^2) = \phi - \frac{e}{2mc^2}A^2 = \text{const} \qquad [15\cdot80]$$

is prepared for equal steps in the value of the *modified potential* $\phi -$

$eA^2/(2mc^2)$. This yields, of course, only the variation of the distance of the electron from the axis and must be supplemented by the graphical or numerical quadrature indicated by Eq. 15·56 yielding the rotation of the electron about the axis:

$$\theta = \theta_o + \int_{z_o}^{z} \frac{1}{r} \left(\frac{eA^2(1 + r'^2)}{2mc^2\phi - eA^2} \right)^{1/2} dz \qquad [15\cdot81]$$

r and r' being taken from the graphical ray trace.

For paraxial rays, which alone enter into consideration for the determination of the cardinal points, focal length, image position, and magnification of the magnetic lens, the numerical methods of section 12·3 may be applied to integrate the differential equation

$$r'' = -\frac{eH^2}{8mc^2\Phi} r \qquad [15\cdot82]$$

assuming the absence of an electric field, and to evaluate the image rotation formula

$$\theta_i - \theta_o = \int_{z_o}^{z_i} \left(\frac{eH^2}{8mc^2\Phi} \right)^{1/2} dz \qquad [15\cdot83]$$

An alternative to this procedure consists in approximating the actual variation of the axial magnetic field by a succession of regions of uniform field and determining the electron path in the substitute field. Plotting

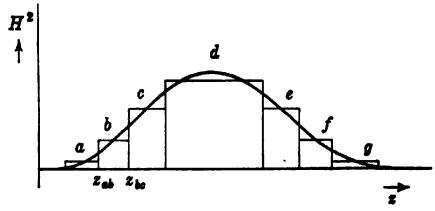


FIG. 15·10. Approximation of Magnetic Field by a Step Function.

H^2 against z (Fig. 15·10), the area under the curve is subdivided into a series of rectangles, each having an area equal to that below the curve segment for the same range of abscissas.

Within any segment, for example, segment b, $H = H_b$ is constant and the general solution of Eq. 15·82 may be written

$$r_b = A \cos\left\{ \left(\frac{eH_b^2}{8mc^2\Phi} \right)^{1/2} (z_b - z_{ab}) \right\} + B \sin\left\{ \left(\frac{eH_b^2}{8mc^2\Phi} \right)^{1/2} (z_b - z_{ab}) \right\}$$

which, with the initial condition $r_b = r_{ab}$ and $r'_b = r'_{ab}$ at $z = z_{ab}$, becomes

$$r_b = r_{ab} \cos\left\{\left(\frac{eH_b^2}{8mc^2\Phi}\right)^{1/2}(z_b - z_{ab})\right\} + r'_{ab}\left(\frac{8mc^2\Phi}{eH_b^2}\right)^{1/2} \sin\left\{\left(\frac{eH_b^2}{8mc^2\Phi}\right)^{1/2}(z_b - z_{ab})\right\} \qquad [15\cdot84]$$

Within a plane rotating about the axis with the electron, the path in the substitute field thus consists of sine-function arcs. These, unlike the parabolic arcs discussed in section 12·4, join together smoothly, since H^2, and hence the second derivative of r in Eq. 15·82, remains finite at the junction between segments. The rotation of the electron about the axis is given by

$$\theta_b - \theta_{ab} = \left(\frac{eH_b^2}{8mc^2\Phi}\right)^{1/2}(z_b - z_{ab}) \qquad [15\cdot85]$$

In the particular case when the magnetic-field distribution is so limited in space that an electron incident parallel to the axis changes its separation from the axis imperceptibly within the field, the magnetic field in question constitutes a " thin " magnetic electron lens. As for a corresponding electric field, the focal length can then be determined by quadrature, since r on the right side of Eq. 15·82 may be treated as a constant. Considering an electron incident on the lens parallel to the axis from object space,

$$\frac{1}{f} = -\frac{r'_i}{r_o} = \frac{e}{8mc^2\Phi}\int_{z_o}^{z_i} H^2\,dz = \frac{0.022}{\Phi}\int_{z_o}^{z_i} H^2\,dz\ \text{cm}^{-1} \qquad [15\cdot86]$$

In the last expression Φ is measured in volts and H in gauss, as usual. The rotation of the image is given, again, by Eq. 15·83.

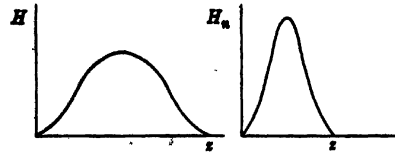


Fig. 15·11. Magnetic Field Distributions with Equal Refractive Power and Different Spatial Extent.

Assume now that the field distribution is contracted by a factor $1/n$ (Fig. 15·11), so that the new distribution is given by

$$H_n(z) = aH(nz) \qquad [15\cdot87]$$

where a is a constant. If

$$\frac{1}{f_n} = \frac{e}{8mc^2\Phi}\int_{z_o}^{z_i} H_n^2(z)\,dz = \frac{1}{f} \qquad [15\cdot88]$$

that is, if the new, thinner lens is to have the same focal length as the original lens, a must be given by $n^{1/2}$. The corresponding rotation of the image becomes, by Eq. 15·83,

$$\begin{aligned}\theta_{ni} - \theta_{no} &= \left(\frac{e}{8mc^2\Phi}\right)^{1/2}\int_{z_o}^{z_i} H_n(z)\,dz \\ &= n^{1/2}\left(\frac{e}{8mc^2\Phi}\right)^{1/2}\int_{z_o}^{z_i} H(nz)\,\frac{d(nz)}{n} = \frac{\theta_i - \theta_o}{n^{1/2}}\end{aligned} \qquad [15\cdot89]$$

It is thus seen that the image rotation of a magnetic lens is steadily reduced as it is made thinner, the refractive power being maintained constant. Consequently, for a very thin lens the image rotation is zero. The image rotation of a magnetic lens approaches π, on the other hand, as the field distribution is made very broad, the uniform field constituting the limiting case (Eq. 15·31).

Magnetic lenses with zero image rotation may be obtained by one other device: If the axial magnetic-field distribution between image and object is made antisymmetric about some point $z = z_a$:

$$H(z) = -H(2z_a - z) \qquad [15\cdot90]$$

the image rotation, as given by Eq. 15·83, vanishes, since it is equal, but for a constant factor, to the area under the field distribution curve (Fig. 15·12a). The refractive power, being proportional to the integral of the square of the axial field strength (Fig. 15·12b), does not vanish unless the field on the axis is equal to zero throughout. An antisymmetric, rotation-free magnetic lens may be realized by using a pair of similar coils traversed by the same current in opposite directions for forming the magnetic field (Fig. 15·12c).

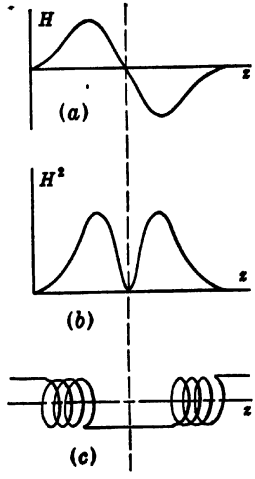


FIG. 15·12. Field Distribution and Coil Arrangement for a Rotation-Free Magnetic Lens.

15·5. Electron Motion in "Two-Dimensional" Magnetic Fields. Two-dimensional magnetic fields, as well as superposed magnetic and

electric fields with other than axial symmetry, are used in a number of practical devices, such as deflection systems, mass spectrographs, cyclotrons and magnetrons. For their consideration it is convenient to return to the equations of motion of the electron (Eq. 15·6) in rectangular coordinates:

$$\ddot{x} = \frac{e}{m}\left(\frac{\partial\phi}{\partial x} - \frac{\dot{y}}{c}h_z + \frac{\dot{z}}{c}h_y\right) \qquad [15\cdot91a]$$

$$\ddot{y} = \frac{e}{m}\left(\frac{\partial\phi}{\partial y} - \frac{\dot{z}}{c}h_x + \frac{\dot{x}}{c}h_z\right) \qquad [15\cdot91b]$$

$$\ddot{z} = \frac{e}{m}\left(\frac{\partial\phi}{\partial z} - \frac{\dot{x}}{c}h_y + \frac{\dot{y}}{c}h_x\right) \qquad [15\cdot91c]$$

Consider, to begin with, an idealized magnetic deflecting field produced, for example, by pole pieces of infinite extent in the z-direction and infinite permeability, whose cross section in any plane normal to the z-axis is sketched, together with the field lines, in Fig. 15·13. As in this case $h_z \equiv 0$ and $\partial h_x/\partial z \equiv \partial h_y/\partial z \equiv 0$, Eqs. 15·91 reduce to

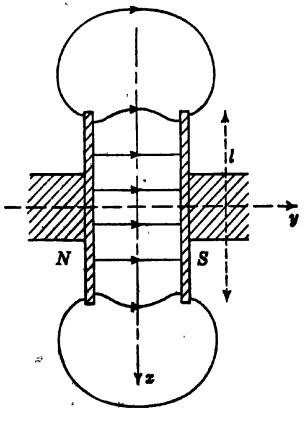


Fig. 15·13. Magnetic Deflecting Field Formed by Two Parallel Flat Pieces.

$$\ddot{x} = \frac{e}{mc}\dot{z}h_y \qquad [15\cdot92a]$$

$$\ddot{y} = -\frac{e}{mc}\dot{z}h_x \qquad [15\cdot92b]$$

$$\ddot{z} = -\frac{e}{mc}(\dot{x}h_y - \dot{y}h_x) \qquad [15\cdot92c]$$

Electrons incident on the deflecting field in the xz-plane of symmetry of the field remain in it, since here h_x and hence $\ddot{y}$ are zero throughout. Both the path and the angle of deflection of an electron incident in this plane can readily be found from Eqs. 15·92a and 15·92c if it is remembered that $z' = dz/dx = \dot{z}/\dot{x}$, $v^2 = \dot{x}^2 + \dot{z}^2$. Thus, by putting $h_y(x,0) = H(x)$,

$$z'' = \frac{\ddot{z}\dot{x} - \dot{x}\ddot{z}}{\dot{x}^3} = -\frac{eH}{mcv}(1 + z'^2)^{3/2} \qquad [15\cdot93]$$

If this is transposed to yield a readily integrable form:

$$\frac{dz'}{(1+z'^2)^{3/2}} = -\frac{eH}{mcv}dx$$

the solution becomes

$$\frac{z'}{(1+z'^2)^{1/2}} = \frac{z'_o}{(1+z'^2_o)^{1/2}} - \int_0^x \frac{eH}{mcv}\,dx \qquad [15\cdot94]$$

If θ is the angle which the ray makes with the x-axis, this may be written in the form

$$\sin\theta = \sin\theta_o - \int_0^x \frac{eH}{mcv}\,dx \qquad [15\cdot95]$$

Assume now that the beam is incident parallel to the axis, that is, that $\theta_o = 0$. Then, if the deflecting field is confined between $x = 0$ and $x = l$, the apparent point of intersection of the deflected ray with the axis is given by

$$x_c = l - \frac{z_l}{\tan\theta_l} = l - \frac{1}{\tan\theta_l}\int_0^l \tan\theta\,dx \qquad [15\cdot96]$$

where $\tan\theta$ is to be regarded as a function of x, as given by Eq. 15·95. The identity of the last two expressions in Eq. 15·96 follows from the fact that $dz = \tan\theta\,dx$.

Equations 15·95 and 15·96 assume familiar forms for a uniform deflecting field of length l, this permitting all the integrations to be carried out in elementary fashion. Thus the angle of deflection is here given by

$$\sin\theta_l = -\frac{eH}{mcv}\,l = -\left(\frac{e}{2mc^2}\right)^{1/2}\frac{Hl}{\phi^{1/2}} = 0.2966\,\frac{Hl}{\phi^{1/2}} \qquad [15\cdot97]$$

and the apparent point of origin of the deflected ray by

$$x_c = l\left(1 - \frac{1}{\sin\theta_l\tan\theta_l}\int_0^{\theta_l}\sin\theta\,d\theta\right) = \frac{\tan\dfrac{\theta_l}{2}}{\sin\theta_l}\,l \qquad [15\cdot98]$$

The first equality follows from Eq. 15·96 if

$$\cos\theta d\theta = -\frac{eH}{mcv}\,dx = \sin\theta_l\,\frac{dx}{l}$$

is substituted. For small angles of deflection the apparent point of origin of the deflected ray on the axis is seen to be located at the center of the field, that is, $x_c = l/2$.

For rays incident parallel to the axis outside the plane of symmetry, the electrons, in addition to being deflected downward in the z-direction, receive a velocity component at right angles to the plane of symmetry equal (by Eq. 15·92b) to

$$\dot{y} = -\int \frac{e}{mc}\,h_z\,dz \qquad [15\cdot99]$$

Since the deflection in the z-direction is, to a first approximation, inde-

pendent of the original separation y_o of the incident ray from the plane of symmetry, and since h_x is, to the same degree of approximation, proportional to y_o (in fact, equal to yH'), Eq. 15·99 indicates that the deflecting field acts simultaneously as a cylindrical lens, causing convergence for a fringe field of the type shown in Fig. 15·13.

The first-approximation differential equation for the deviation in the y-direction can be derived from Eqs. 15·92a and 15·92b in similar manner as Eq. 15·93:

$$y'' = -\frac{e}{mcv}\frac{\tan^2\theta}{\sin\theta}(H'y + Hy') \qquad [15\cdot100]$$

where $\sin\theta$ and $\tan\theta$ are given by Eq. 15·95 with $\theta_o = 0$. This equation may, for any given field variation in the plane of symmetry, be integrated by the usual methods of numerical integration (section 12·3).

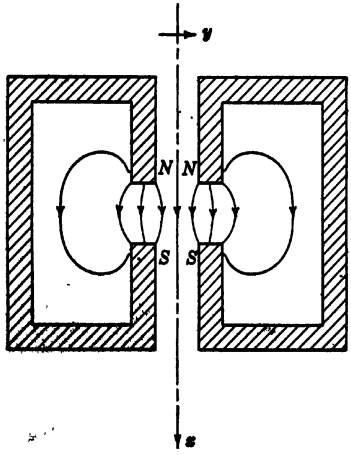


Fig. 15·14. A Magnetic Cylinder Lens.

For a more detailed discussion of the effects produced by the non-homogeneous character of the deflecting field the reader is referred to the treatment given by Maloff and Epstein.[5]

The magnetic deflecting field is not, however, the true analogue of the electric cylinder lens.[6] In a magnetic cylinder lens the field in the plane of symmetry is $h_x(x,0) = H(x)$, that is, the field vector lies in the plane of symmetry and is normal to the generatrices of the cylindrical surfaces of the configuration, as shown in Fig. 15·14. Here the field is shown generated by two identical cylindrical permanent magnets. More commonly, it would be produced by two pairs of cylindrical pole pieces of an electromagnet.

For such a field it is possible to write, in view of Eq. 11·55 for an analogous electric field,

$$\psi = \Psi - \frac{y^2}{2}\Psi'' + \cdots \qquad h_y = -yH' + \cdots$$
$$h_z = H - \frac{y^2}{2}H'' + \cdots \qquad [15\cdot101]$$

[5] See reference 5.
[6] See section 11·5.

where $\psi(x,y)$ is the magnetic scalar potential and $\Psi(x)$ its value in the xz-plane. Thus Eqs. 15·92 become, to a first approximation,

$$\ddot{x} = -\frac{e}{mc}\dot{z}yH' \qquad [15\cdot102a]$$

$$\ddot{y} = -\frac{e}{mc}\dot{z}H \qquad [15\cdot102b]$$

$$\ddot{z} = \frac{e}{mc}(\dot{x}yH' + \dot{y}H) \qquad [15\cdot102c]$$

The corresponding path equations may be derived as before:

$$y'' = \frac{\ddot{y}\dot{x} - \ddot{x}\dot{y}}{\dot{x}^3} = -\left(\frac{e}{2mc^2\Phi}\right)^{1/2} z'H \qquad [15\cdot103a]$$

$$z'' = \frac{\ddot{z}\dot{x} - \ddot{x}\dot{z}}{\dot{x}^3} = \left(\frac{e}{2mc^2\Phi}\right)^{1/2}(yH' + y'H) \qquad [15\cdot103b]$$

In the approximation here employed $\dot{x} = (2e\Phi/m)^{1/2}$.

To gain insight into the action of such a magnetic field, assume that the field is " short " and consider rays incident parallel and close to the x-axis in the xy-plane. Equation 15·103b may be integrated directly, yielding

$$z' = \left(\frac{e}{2mc^2\Phi}\right)^{1/2} yH \qquad [15\cdot104]$$

which, substituted in Eq. 15·103a, gives

$$y'' = \frac{-e}{2mc^2\Phi} yH^2 \qquad [15\cdot105]$$

or, for $y = y_o$ in the region where $H \neq 0$ (that is, for a thin lens),

$$y' = -\frac{y_o e}{2mc^2\Phi}\int H^2\,dx \qquad [15\cdot106]$$

Hence all paraxial rays incident parallel to the x-axis intersect the plane of symmetry at a distance from the lens field $x = f$:

$$\frac{1}{f} = -\frac{y'}{y_o} = \frac{e}{2mc^2\Phi}\int H^2\,dx = \frac{0.088}{\Phi}\int H^2\,dx \text{ cm}^{-1} \qquad [15\cdot107]$$

which is just four times the refractive power of the axially symmetric thin lens with the same axial field distribution (Eq. 15·86). However, the rays incident parallel to the axis in the xy-plane do not come together in a point, as for the electric cylindrical lens, but strike the vertical

line $x = f$, $y = 0$ at a distance (Eq. 15·104)

$$z_f = y_o \left(\frac{e}{2mc^2\Phi}\right)^{1/4} \int H\, dx \qquad [15\cdot108]$$

proportional to the original distance y_o of the ray from the axis, above the xy-plane, as shown in Fig. 15·15. The conditions for the vanishing of this vertical displacement are the same as the conditions for the vanishing of the image rotation in the case of axially symmetric magnetic lenses, discussed in the preceding section.

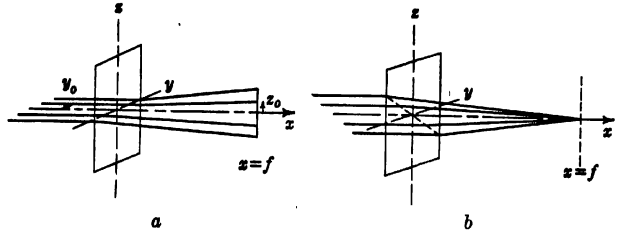


FIG. 15·15. Action of a Magnetic Cylinder Lens on a Flat Parallel Horizontal (*a*) and Suitably Inclined Pencil (*b*).

15·6. Superposed Electric and Magnetic Fields. As early as 1897 J. J. Thomson[7] utilized the characteristics of the motion of electric charges in crossed uniform electric and magnetic fields to measure the velocity of cathode rays, preliminary to the determination of the specific charge of the electron. Since then crossed fields have been employed in

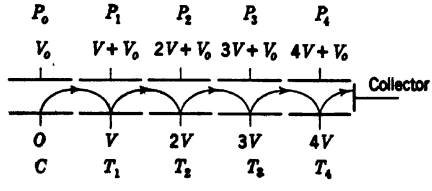


FIG. 15·16. Electrode Arrangement and Electron Paths in Magnetic Electron Multiplier (Schematic).

numerous practical electronic devices. Thus the condition of a crossed uniform electrostatic and uniform magnetic field is approximated in the magnetic electron multiplier (Fig. 15·16),[8] where electrons leaving a photocathode C and succeeding secondary-emission targets T_1, T_2, $\cdots$

[7] See reference 6.

[8] See Zworykin, Morton, and Malter, reference 7.

are accelerated by the electric field between these electrodes and the field plates P_o, P_1, P_2, $\cdots$ opposite the targets, but are turned back to the next target electrode by the superposed uniform magnetic field normal to the plane of the sketch. A similar field combination is employed for the horizontal deflection in the Orthicon,[9] a television pickup tube.

With the magnetic field in the z-direction and the uniform electric field $E = -\partial\phi/\partial y$ in the y-direction, Eqs. 15·91 become

$$\ddot{x} = -\frac{e}{mc}\dot{y}H \qquad [15\cdot109a]$$

$$\ddot{y} = -\frac{e}{mc}(cE - \dot{x}H) \qquad [15\cdot109b]$$

$$\ddot{z} = 0 \qquad [15\cdot109c]$$

All these equations can be integrated. For an electron starting from the point (0,0,0) with the velocity components $\dot{x}_o$, $\dot{y}_o$, $\dot{z}_o$:

$$\dot{x} - \dot{x}_o = -\frac{e}{mc}yH \qquad [15\cdot110a]$$

$$\dot{y} - \dot{y}_o = -\frac{e}{mc}(cEt - xH) \qquad [15\cdot110b]$$

$$\dot{z} = \dot{z}_o t \qquad [15\cdot110c]$$

Substituting the first two differential equations in Eqs. 15·109a and 15·109b leads to the two ordinary differential equations

$$\ddot{x} = -\left(\frac{eH}{mc}\right)^2\left(x + \frac{mc\dot{y}_o}{eH}\right) + \left(\frac{e}{m}\right)^2\frac{EH}{c}t \qquad [15\cdot111a]$$

$$\ddot{y} = -\left(\frac{eH}{mc}\right)^2\left(y - \frac{mc\dot{x}_o}{eH} + \frac{mc^2E}{eH^2}\right) \qquad [15\cdot111b]$$

These are linear differential equations with constant coefficients and can be solved by standard methods,[10] yielding

$$x = \frac{cEt}{H} - \frac{mc\dot{y}_o}{eH} + \frac{mc}{eH}\left[\left(\dot{x}_o - \frac{cE}{H}\right)^2 + \dot{y}_o^2\right]^{1/2}\sin\left(\frac{eHt}{mc} - \arctan\frac{\dot{y}_o}{\frac{cE}{H} - \dot{x}_o}\right) \qquad [15\cdot112a]$$

[9] See Rose and Iams, reference 8.

[10] See Adams, reference 9, Formulae 8.100–8.111.

$$y = -\frac{mc}{eH}\left(\frac{cE}{H} - \dot{x}_o\right)$$
$$-\frac{mc}{eH}\left[\left(\dot{x}_o - \frac{cE}{H}\right)^2 + \dot{y}_o^2\right]^{1/2} \cos\left(\frac{eHt}{mc} - \arctan\frac{\dot{y}_o}{\frac{cE}{H} - \dot{x}_o}\right) \qquad [15\cdot112b]$$

Some of the paths represented by these parametric equations (with $\dot{z}_o = 0$) are shown in Fig. 15·17. For zero initial velocity they are simple cycloids, corresponding to the path described by a point P on

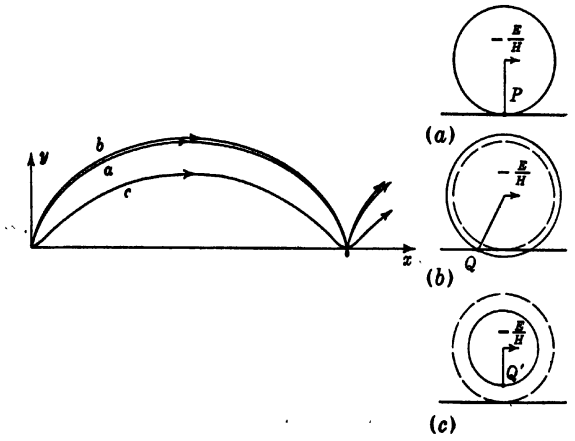


FIG. 15·17. Paths of Electrons in Crossed Uniform Fields for Zero Initial Velocity (*a*), an Initial Velocity in the y-Direction (*b*), and an Initial Velocity in the x-Direction (*c*).

the rim of a wheel rolling on the x-axis with a forward velocity $10^8 \cdot E/H$ centimeters per second and a radius $R = -mc^2E/(eH^2) = -5.68E/H^2$ centimeters, with E measured in volts per centimeter and H measured in gauss. The distance between successive arrivals at the x-axis is

$$d = 2\pi R = \frac{35.7E}{H^2} \text{ cm} \qquad [15\cdot113]$$

If the electron enters the field with a certain initial velocity $\dot{x}_o$, y_o, the path is such as is described by a point Q or Q' on a spoke of the wheel a distance

$$\frac{mc}{eH}\left[\left(\dot{x}_o - \frac{cE}{H}\right)^2 + \dot{y}_o^2\right]^{1/2}$$

from the center. This may be either greater or less than R, giving rise to a *prolate* or *curtate* cycloid (curves b and c of Fig. 15·17). Although the velocity of the wheel is unchanged, the level of the "track" on which the wheel rolls is raised (or lowered) by an amount $mc\dot{x}_o/(eH)$.

The compound field considered may also be used to deflect electrons entering the field with a considerable velocity in the z-direction, that is, normal to both fields. In this case Eqs. 15·112 give the path of the electron in the crossed uniform fields if t is replaced by $z/\dot{z}_o$. Figure 15·18a and b shows the projection of this path on the xz- and yz-planes,

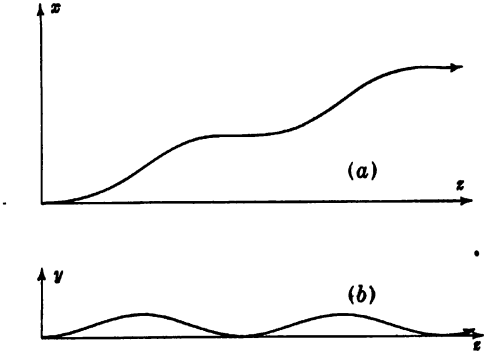


FIG. 15·18. Superposed Uniform Electric Field in y-Direction and Uniform Magnetic Field in z-Direction as Deflecting Fields. Projection of Electron Paths on the xz- and yz-Planes.

respectively. It is seen that the action of this deflecting field is quite different from that of the simple magnetic or electric deflecting field. It does not produce a change in direction of the incident beam, but merely a displacement in the x-direction, provided that the incident velocity of the beam and the magnetic field are such that an integral number (n) of loops are executed during the traversal of the crossed fields:

$$\frac{eHt}{mc} = \frac{eHl}{mc\dot{z}_o} = \left(\frac{e}{2mc^2}\right)^{\frac{1}{2}} \frac{Hl}{\phi_o^{\frac{1}{2}}} = 2\pi n \qquad [15·114]$$

where l is the length of the deflecting field and ϕ_o is the potential corresponding to the velocity of the incident electrons. The required value of the magnetic field is therefore

$$H = \frac{2\pi cn\left(\frac{2m}{e}\right)^{\frac{1}{2}}\phi_o^{\frac{1}{2}}}{l} = \frac{21.2n\phi_o^{\frac{1}{2}}}{l} \text{ gauss} \qquad [15·115]$$

The deflection in the x-direction is proportional to the electric field E and given by nd, where d has the value indicated by Eq. 15·113.

The loops characteristic of electrons entering crossed uniform fields with zero initial velocity normal to the magnetic field are disadvantageous for deflection systems of this type, since they give rise to a relatively large spreading of the beam if the beam contains electrons of slightly varying velocity, so that not all the electrons can simultaneously satisfy the condition 15·114. Furthermore, many applications require the electrons to enter a space free from transverse electric field after deflection, a transition which cannot be carried out abruptly. Hence the electric deflecting fields normally employed will not be uniform, but, preferably, will taper off at both ends. The design of suitable deflecting plates presents an interesting example of electron-optical method.

An application of Laplace's equation to the field of the given symmetry makes it possible to express the potential variation in the form

$$\phi = \Phi_o - yE + \frac{y^3}{6}E'' - \frac{y^5}{120}E^{\mathrm{IV}} + \cdots \qquad [15\cdot116]$$

where $E(z)$ is the electric field in the plane of symmetry, directed in the y-direction.

Whereas Eqs. 15·109a and 15·109b apply, to a first approximation, also here, Eq. 15·109c should be replaced by

$$\ddot{z} = \frac{e}{m}\frac{\partial\phi}{\partial z} = -\frac{e}{m}yE' \qquad [15\cdot117]$$

The term on the right is small, however, and, at least for moderate deflections, it may be neglected, leading again to Eq. 15·110c with $\dot{z}_o = (2e\Phi_o/m)^{1/2}$. If this relation is used to eliminate the time parameter in Eqs. 15·109a and 15·109b, there results, with zero initial velocities assumed in the x- and y-directions:

$$x' = -\frac{e}{mc\dot{z}_o}yH = -\left(\frac{eH^2}{2mc^2\Phi_o}\right)^{1/2}y \qquad [15\cdot118a]$$

$$y'' = -\frac{E}{2\Phi_o} + \left(\frac{eH^2}{2mc^2\Phi_o}\right)^{1/2}x' = -\frac{E}{2\Phi_o} - \frac{eH^2}{2mc^2\Phi_o}y \qquad [15\cdot118b]$$

With the aid of these two equations it is possible to determine the field variation $E(z)$ resulting in a prescribed path $x(z)$, $y(z)$:

$$E = -\frac{eH^2}{mc^2}y - 2\Phi_o y'' \qquad [15\cdot119]$$

For example, let the prescribed path be

$$x = A\left(\arctan z + \frac{\pi}{2}\right) \qquad [15\cdot120]$$

As shown in Fig. 15·19*a*, the electron enters the field parallel to the axis and leaves it, again in the xz-plane, in the same direction, a distance $A\pi$ from the axis. The constant A is supposed to be made smaller than

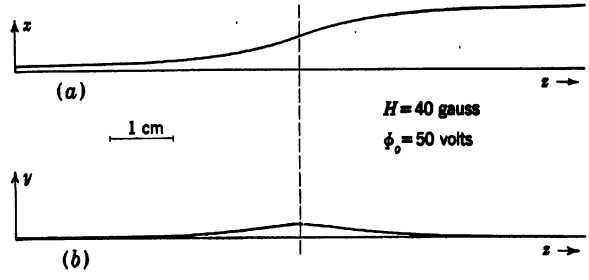


FIG. 15.19. Required Path of Electron. Projection on xz- and yz-Planes.

unity to justify the approximations in the present calculation. The motion in a direction normal to the xz-plane is given by Eq. 15·118*a*:

$$y = -\left(\frac{2mc^2\Phi_0}{eH^2}\right)^{1/2} \frac{A}{1+z^2} = -3.37\,\frac{\Phi_0^{1/2}}{H}\,\frac{A}{1+z^2}\ \text{cm} \quad [15\cdot121]$$

That is, it consists of a single loop with a maximum displacement $-3.37A\Phi_0^{1/2}/H$ at the center of the field.

Equation 15·119, determining the field variation in the plane of symmetry, becomes

$$E = 0.593H\Phi_0^{1/2}\frac{A}{1+z^2} + 13.48\,\frac{\Phi_0^{3/2}}{H}\,A\,\frac{3z^2-1}{(1+z^2)^3}\ \text{volt per cm} \quad [15\cdot122]$$

With Eq. 15·122 substituted in Eq. 15·116, the function $(\phi - \Phi_0)/A$, that is, the potential distribution for any given value of A (or strength of the deflecting field E) is given. Ascribing a fixed value to this function, Eq. 15·116 becomes the equation of the intersection of an equipotential surface with the xz-plane. In order to obtain a deflecting field of the desired variation, it is necessary to shape the deflecting plates in the form of a pair of these equipotential surfaces. Now the determination of these shapes involves the solution of the equation

$$y\frac{E(z)}{A} - \frac{y^3}{6}\frac{E''(z)}{A} + \frac{y^5}{120}\frac{E^{IV}(z)}{A} - \cdots = \frac{\Phi_0 - \phi}{A} = \text{const} \quad [15\cdot123]$$

for y for given values of z, a procedure which becomes very slow if more than the first two terms on the left are considered. For this reason, as well as for minimizing the effect of external disturbing fields and fringe fields, it is normally desirable to place the plates as close together as

possible. The limit for this is given by the deflection in the y-direction, whose maximum value is given by Eq. 15·121:

$$y_{max} = -3.37 \frac{\Phi_o^{1/2}}{H} A_{max} = -1.07 \frac{\Phi_o^{1/2}}{H} x_{max} \text{ cm} \qquad [15·124]$$

where x_{max} is the maximum total deflection in the x-direction required. Assume, for example,

$$x_{max} = 1 \text{ cm} \qquad \Phi_o = 50 \text{ volts} \qquad H = 40 \text{ gauss}$$

then $$A_{max} = 0.318 \text{ cm} \qquad y_{max} = 0.2 \text{ cm}$$

If it should be assumed that the approach of the electron paths to the deflecting plates is closest at the center of the field, substitution of the above values of Φ_o, H, and y in Eqs. 15·122 and 15·123 with $z = 0$ will yield the value of the constant

$$\frac{\Phi_o - \phi}{A} = 8.2 \text{ volts per cm}$$

for that equipotential surface which is just reached by the electron for the maximum deflection in the x-direction. If the deflecting plates are shaped to follow this equipotential (and that for the value of the constant -8.2), the voltage which has to be applied to them for the required maximum deflection will be

$$2(\Phi_o - \phi) = 5.2 \text{ volts}$$

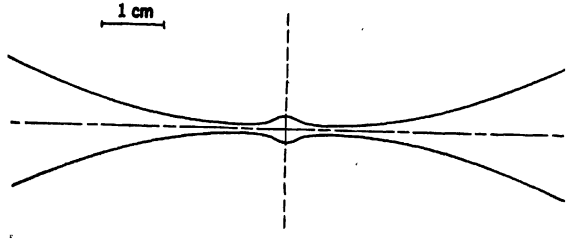


Fig. 15·20. Shape of Deflecting Plates Resulting in Path of Fig. 15·19.

Evaluating the curve for the equipotentials, by solving Eq. 15·123 for y, neglecting the third term on the left, for the prescribed value of the constant involves the solution of a cubic equation.[11] The result is shown in Fig. 15·20. It is seen that the plates approach the path closest not at the center, but about 0.5 centimeter away from the center. In fact the maximum deflection in the x-direction possible with these plates is less than half that originally prescribed.

[11] See Adams, reference 9, Formula 1.271.

A simple arrangement for demonstrating the motion of electrons in crossed electric and magnetic fields has been devised by A. Rose of the RCA Laboratories.[12] Consider (Fig. 15·21) a long, thin permanent mag-

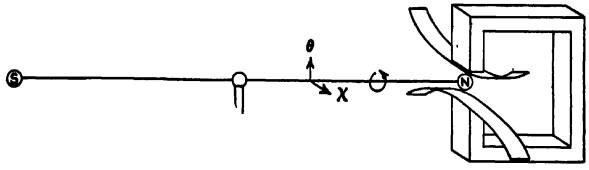


FIG. 15·21. Mechanical Model for Determining Electron Paths in Crossed Fields.

net, of length $2R$, rotating at a high angular velocity ν about its axis and supported in a frictionless universal joint at its center of gravity. When one end of the magnet with pole-strength μ is subjected to a magnetic field H, the motion of the magnet will be given by[13]

$$A\ddot{\theta} + C\nu\dot{\chi} = \mu H_\theta R \qquad [15\cdot125a]$$

$$A\ddot{\chi} - C\nu\dot{\theta} = \mu H_\chi R \qquad [15\cdot125b]$$

It is assumed that the angular deviations θ, χ of the axis of the magnet from its original position remain small. H_θ and H_χ are the components of H in the θ and the χ-directions. C and A are the moments of inertia of the magnet about its axis of rotation and an axis through the center of gravity normal to the axis of rotation, respectively. Since, with the assumptions of the smallness of θ and χ here made, it is permissible to write

$$R\theta = y \qquad R\chi = x \qquad [15\cdot126]$$

Equation 15·125 may also be written in the form

$$\ddot{y} = \frac{\mu R^2}{A} H_y - \frac{C\nu}{A}\dot{x} \qquad [15\cdot127a]$$

$$\ddot{x} = \frac{\mu R^2}{A} H_x + \frac{C\nu}{A}\dot{y} \qquad [15\cdot127b]$$

These two equations are seen to pass over into Eqs. 15·91a and 15·91b for the case $h_x = h_y = 0$ if the following substitutions are made:

$$\frac{\mu R^2}{A} H_y \to -\frac{e}{m} E_y \qquad \frac{\mu R^2}{A} H_x \to -\frac{e}{m} E_x$$

$$\frac{C\nu}{A} \to -\frac{eh_z}{mc} \qquad [15\cdot128]$$

[12] See Rose, reference 10.

[13] See Osgood, reference 11, p. 235, Problem 3.

Under certain conditions the motion of the rotating magnet pole in the magnetic field thus becomes identical with the motion of an electron in an electric field similar to the magnetic field. These are that a uniform magnetic field of the same direction as, and proportional to, the spin of the magnet is superposed on the electric field and that the velocity of the electron in the direction of the uniform magnetic field is zero. If the velocity of the electron in the direction of the magnetic field is not zero, the analogy still holds for the projection of the path of the electron on a plane normal to the magnetic field (the xy-plane), provided that the variation of the magnetic field in the model is made to correspond to the variation of the electric field along the path of the electron. Thus the first substitution in Eq. 15·128 might be written

$$\frac{\mu R^2}{A} H_y(x(z)) \rightarrow -\frac{e}{m} E_y(z)$$

if the y-component of the electric field may be regarded as a function of z only. Here $x(z)$ is the x-component of the path. Since this is normally not known, being in fact part of the desired information, it is desirable to replace it by some readily determinable approximation, for which

$$x(z) = \frac{1}{\dot{z}_o} \int \frac{cE_y}{H} dz \qquad [15·129]$$

may serve.

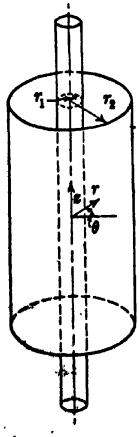


Fig. 15·22. Electrode Arrangement in Magnetron.

The mechanical model demonstrates strikingly the reduction in the amplitude of the loops executed by the electron if the electron is permitted to enter a gradually increasing electric deflecting field instead of one which reaches its maximum value almost instantly.

Two other types of superposed electric and magnetic fields will be considered briefly: A crossed uniform magnetic and radial electric field, the basic configuration of the magnetron,[14] and a superposed cylindrical magnetic and electric field. In the first case consider (Fig. 15·22) two coaxial cylinders, the inner of radius r_1 and potential V_1 (the cathode) and the outer of radius r_2 and potential V_2 (the anode). The electric potential distribution between the cylinders is then given

[14] See Hull, reference 12.

according to Eq. 11·24 by

$$\phi(r) = (V_2 - V_1)\frac{\log \dfrac{r}{r_1}}{\log \dfrac{r_2}{r_1}} + V_1 \qquad [15\cdot130]$$

Let a uniform magnetic field H parallel to the axis (the z-direction) be superposed on the electric field. The equations of motion of an electron in this compound field are then (Eq. 15·6):

$$\ddot{r} - r\dot{\theta}^2 = -\frac{e}{mc}Hr\dot{\theta} + \frac{e}{m}\frac{V_2 - V_1}{\log \dfrac{r_2}{r_1}}\frac{1}{r} \qquad [15\cdot131a]$$

$$\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta}) = \frac{e}{mc}H\dot{r} \qquad [15\cdot131b]$$

$$\ddot{z} = 0 \qquad [15\cdot131c]$$

Consider an electron leaving the point $r = r_1$, $\theta = z = 0$ on the inner cylinder with zero velocity at the time $t = 0$. The integration of Eq. 15·131b, multiplied through by r, yields

$$\dot{\theta} = \frac{eH}{2mc}\left(1 - \frac{r_1^2}{r^2}\right) \qquad [15\cdot132]$$

The electron leaves the surface of the inner cylinder perpendicularly to the surface and is deflected in a counterclockwise direction by the magnetic field. Unless it is intercepted by the anode, it will attain a maximum separation from the axis, r_{max}, and then return toward the cathode. At the point of greatest distance from the axis the electron moves in a tangential direction so that all its kinetic energy is given by $(m/2)r_{max}^2\dot{\theta}^2$:

$$\frac{mr_{max}^2\dot{\theta}^2}{2} = \frac{m}{2}\left(\frac{eH}{2mc}\right)^2\frac{(r_{max}^2 - r_1^2)^2}{r_{max}^2} = \frac{e(V_2 - V_1)\log \dfrac{r_{max}}{r_1}}{\log \dfrac{r_2}{r_1}} \qquad [15\cdot133]$$

For $r_2 = r_{max}$, that is, the condition in which the electron barely reaches the anode, the relation between the magnetic field, the anode voltage, and the electrode dimensions becomes

$$H = \left(\frac{8mc^2(V_2 - V_1)}{e}\right)^{1/2}\frac{r_2}{r_2^2 - r_1^2} = 6.75V^{1/2}\frac{r_2}{r_2^2 - r_1^2}\ \text{gauss} \qquad [15\cdot134]$$

where $V = V_2 - V_1$. The actual path of the electron may be obtained by substituting Eq. 15·132 in Eq. 15·131a:

$$\ddot{r} = -\left(\frac{eH}{2mc}\right)^2 \frac{r^4 - r_1^4}{r^3} + \frac{e(V_2 - V_1)}{mr \log \dfrac{r_2}{r_1}} \qquad [15\cdot135]$$

Multiplying the above by $\dot{r}\,dt$ and integrating leads to

$$\dot{r}^2 = 2(V_2 - V_1)\frac{e}{m}\frac{\log \dfrac{r}{r_1}}{\log \dfrac{r_2}{r_1}} - \left(\frac{eH}{2mc}\right)^2 \frac{(r^2 - r_1^2)^2}{r^2} \qquad [15\cdot136]$$

This, finally, can be translated into a simple quadrature by taking the square root, dividing by the right side of the equation, and multiplying by dt:

$$t = \int_{r_1}^{r} \frac{dr}{\left\{2(V_2 - V_1)\dfrac{e}{m}\dfrac{\log \dfrac{r}{r_1}}{\log \dfrac{r_2}{r_1}} - \left(\dfrac{eH}{2mc}\right)^2 \dfrac{(r^2 - r_1^2)^2}{r^2}\right\}^{1/2}} \qquad [15\cdot137]$$

For any specific values of the parameters r_1, r_2, H, and $V_2 - V_1$ the function $t(r)$ (and hence the function $r(t)$) can be calculated numerically from $r = r_1$ to $r = r_{max}$. Beyond the latter point the square root under the integral sign changes sign. The function $r(t)$ is symmetrical about the point $t_m = t(r_{max})$ in the sense that

$$r(t) = r(2t_m - t)$$

The path becomes fully determined in parametric form after the additional integration of Eq. 15·132:

$$\theta = \int_0^t \frac{eH}{2mc}\frac{r^2 - r_1^2}{r^2}\,dt \qquad [15\cdot138]$$

Figure 15·23 shows the path of an electron when it just reaches the anode ($r_{max} = r_2$) for $r_1 = r_2/4$ and for $r_1 \cong 0$. In the latter case the path becomes a circle, since the electrons travel, except in the immediate neighborhood of the cathode, with their full maximum velocity in a uniform magnetic field. In most practical cases it is necessary to take account of the effect of space charge on the paths of the electrons. Its primary effect is to lower the potential near the cathode, since the space

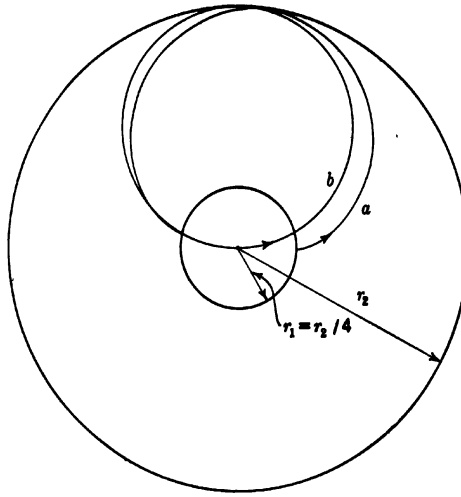


FIG. 15·23. Electron Paths in Space-Charge-Free Magnetron: (*a*) for $r_1 = r_2/4$, (*b*) for $r_1 \doteq 0$.

charge is densest there. This causes a deviation from the circular path even for the case $r_1 \doteq 0$. For details regarding the effect of space charge on the paths the reader is referred to Hull's original article.

The second field distribution, that is, two superposed cylindrical magnetic and electric fields, may in principle be realized by two coplanar semi-infinite electrodes at different potentials ϕ_1 and ϕ_2 meeting at the z-axis placed over two similar infinitely permeable pole pieces of a permanent or an electromagnet (Fig. 15·24).[15] This type of field is of interest primarily because, like the uniform magnetic field superposed on a uniform electric field, it can be employed for image formation.

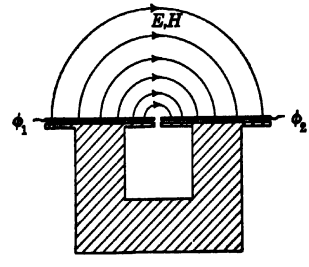


FIG. 15·24. Production of Superposed Cylindrical Electric and Magnetic Fields.

For the configuration described the electric and magnetic fields are given by[16]

$$E_\theta = -\frac{1}{r}\frac{\partial \phi}{\partial \theta} = \frac{\phi_1 - \phi_2}{\pi r} \qquad H = -\frac{1}{r}\frac{\partial \psi}{\partial \theta} = \frac{\psi_1 - \psi_2}{\pi r} \qquad [15\cdot139]$$

[15] See Rose, reference 13.
[16] See Eq. 11·28.

ψ being the magnetic scalar potential. The field lines of both fields are circles about the z-axis. Omitting the subscript θ and writing E_o, H_o for the field strengths along circles of unit radius about the z-axis,

$$E = \frac{E_o}{r} \qquad H = \frac{H_o}{r} \tag{15·140}$$

The equations of motion in this field become then

$$\ddot{r} - r\dot{\theta}^2 = \frac{e}{mc}\frac{H_o\dot{z}}{r} \tag{15·141a}$$

$$\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta}) = -\frac{e}{m}\frac{E_o}{r} \tag{15·141b}$$

$$\ddot{z} = -\frac{e}{mc}\frac{H_o\dot{r}}{r} \tag{15·141c}$$

The last two equations can immediately be integrated, leading to

$$r^2\dot{\theta} - r_o^2\dot{\theta}_o = -\frac{e}{m}E_o t \tag{15·142a}$$

$$\dot{z} - \dot{z}_o = -\frac{eH_o}{mc}\log\frac{r}{r_o} \tag{15·142b}$$

This set of equations is not readily integrated exactly. Hence two important limiting cases will be considered separately: (1) that of infinitesimal initial velocity components $r_o\dot{\theta}_o$, $\dot{z}_o$, $\dot{r}_o$, and non-vanishing electric field and (2) that of a finite initial velocity component $r_o\dot{\theta}_o$, infinitesimal components $\dot{z}_o$, $\dot{r}_o$, and zero electric field. In both cases it will be assumed that

$$\frac{eH_o^2}{8mc^2\phi_2} \gg 1 \tag{15·143}$$

This signifies that, if H_o were a uniform magnetic field, the electron considered would execute several complete loops about the field direction. It may be regarded quite generally as the condition for which the electron follows the magnetic field lines in its course. It will follow them more closely, spiraling about them in a tighter helix, in proportion as H^2/ϕ is larger.

With the above assumptions $r - r_o$ may also be regarded as small and Eq. 15·142a may be integrated to yield, as a first approximation,

$$\theta - \theta_o = \dot{\theta}_o t - \frac{eE_o}{2mr_o^2}t^2 \quad \text{or} \quad t = r_o\left(\frac{2m(\theta-\theta_o)}{-eE_o}\right)^{1/2} + \frac{mr_o^2}{eE_o}\dot{\theta}_o \tag{·144}$$

Furthermore, Eq. 15·141*a*, with Eq. 15·142*a* substituted in it, takes on the approximate form

$$\dot{r} = r_o\left(\frac{r_o^2\dot{\theta}_o - \dfrac{eE_o t}{m}}{r_o^2}\right)^2 + \frac{eH_o}{mcr_o}\left(\dot{z}_o - \frac{eH_o}{mc}\frac{r - r_o}{r_o}\right) \qquad [15\cdot145]$$

This may be integrated by standard methods[17] to yield

$$r - r_o = -2r_o\left(\frac{mc^2E_o}{eH_o^2}\right)^2\left(1 - \cos\Omega - \frac{\Omega^2}{2}\right) + \dot{r}_o\frac{mcr_o}{eH_o}\sin\Omega + \dot{z}_o\frac{mcr_o}{eH_o}(1 - \cos\Omega) \qquad [15\cdot146]$$

with

$$\Omega = \left[\frac{2eH_o^2(\theta - \theta_o)}{-mc^2E_o}\right]^{1/2} = \frac{eH_o}{mcr_o}\left(t - \frac{mr_o^2}{eE_o}\dot{\theta}_o\right) \qquad [15\cdot147]$$

By utilizing Eq. 15·146, Eq. 15·142*b* is solved by

$$z - z_o = 2r_o\left(\frac{mc^2E_o}{eH_o^2}\right)^2\left(\Omega - \sin\Omega - \frac{\Omega^3}{6}\right) + \dot{z}_o\frac{mcr_o}{eH_o}\sin\Omega - \dot{r}_o\frac{mcr_o}{eH_o}(1 - \cos\Omega) \qquad [15\cdot148]$$

It is seen that for $\Omega = 2n\pi$ all the terms involving the initial velocities vanish, indicating that in the corresponding planes

$$\theta = \theta_o - \frac{(2n\pi c)^2mE_o}{2eH_o^2} \qquad [15\cdot149]$$

true images of the object are formed. The remaining terms indicate that these images are not free from distortion. As the distortion terms are proportional to r_o, the image shift is seen to increase linearly outward, taking the form of a simultaneous stretch in the r-direction and a shift in the z-direction, as shown in Fig. 15·25. Figure 2·19 shows an actual electron picture obtained with a field configuration of the type discussed. The original "resolution pattern" was projected on a photocathode forming one of the two flat electrodes.

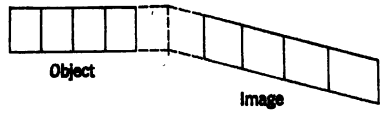


FIG. 15·25. Stretch and Shear Distortion in Image Formed by Superposed Electric and Magnetic Fields.

[17] See Adams, reference 9, Formula 8.410.

It is interesting to study the variation in image distortion with the order of the image, if the position of the latter is kept fixed. Thus if, for example, $\theta - \theta_o = \pi$, and hence, by Eq. 15·149,

$$\frac{mc^2E_o}{eH_o^2} = -\frac{1}{2\pi n^2} \qquad [15\cdot150]$$

the position of the image points is given by

$$r - r_o = \frac{(2\pi n)^2}{(2\pi n^2)^2} r_o = \frac{r_o}{n^2} \qquad [15\cdot151a]$$

$$z - z_o = \frac{2}{(2\pi n^2)^2}\left(2\pi n - \frac{(2\pi n)^3}{6}\right) r_o \approx -\frac{2\pi}{3n} r_o = -\frac{2.09}{n} r_o \qquad [15\cdot151b]$$

Thus both distortions decrease rapidly with the order of the image or the strength of the magnetic field.

In the second case, that is, electrons of uniform velocity $r_o\dot{\theta}_o = (2e\phi_o/m)^{\frac{1}{2}}$ traveling in a cylindrical magnetic field with infinitesimal lateral initial velocity components $\dot{r}_o$ and $\dot{z}_o$, Eqs. 15·142 reduce to

$$\dot{\theta} = \left(\frac{r_o}{r}\right)^2 \dot{\theta}_o \qquad [15\cdot152a]$$

$$\dot{z} = \dot{z}_o - \frac{eH_o}{mc} \log \frac{r}{r_o} \qquad [15\cdot152b]$$

If, with consistent neglect of small quantities, this is substituted in Eq. 15·141*a*, it follows that

$$\ddot{r} = r_o\dot{\theta}_o^2\left(1 - 3\frac{r - r_o}{r_o}\right) + \frac{eH_o}{mcr_o}\left(\dot{z}_o - \frac{eH_o}{mcr_o}(r - r_o)\right) \qquad [15\cdot153]$$

By utilising the inequality (15·143) and setting

$$\Omega = \left(\frac{eH_o^2}{2mc^2\phi_o}\right)^{\frac{1}{4}} (\theta - \theta_o) \cong \frac{eH_o t}{mcr_o} \qquad [15\cdot154]$$

Eqs. 15·153 and 15·152 are solved approximately by

$$r - r_o = \frac{mcr_o}{eH_o}\left(\dot{r}_o - \frac{c\Delta\phi}{H_o}\Omega\right)\sin\Omega + \frac{mcr_o}{eH_o}\left(\dot{z}_o + \frac{2c\phi_o}{H_o} + \frac{2c\Delta\phi}{H_o}\right)(1 - \cos\Omega) \qquad [15\cdot155a]$$

$$z - z_o = -\frac{2mc^2\phi_o}{eH_o^2} r_o\Omega - \frac{mcr_o}{eH_o}\left(\dot{r}_o + \frac{c\Delta\phi}{H_o}\Omega\right)(1 - \cos\Omega) + \frac{mcr_o}{eH_o}\left(\dot{z}_o + \frac{2c\phi_o}{H_o} + \frac{2c\Delta\phi}{H_o}\right)\sin\Omega - \frac{mc^2\Delta\phi}{eH_o^2} r_o\Omega \qquad [15\cdot155b]$$

Here it has been assumed that, in addition to the initial velocities in the r- and z-directions, there may be a variation $e\Delta\phi$, also infinitesimal in character, in the original kinetic energy associated with motion in the θ-direction. It is seen that also in this case an image is formed for

$$\Omega = 2n\pi$$

The only distortion term is a linear shift, or "shear," in the z-direction which, for $\theta - \theta_o = \pi$, becomes

$$z - z_o = -r_o \frac{2n\pi}{(2n)^2} = -\frac{1.57 r_o}{n} \qquad [15\cdot156]$$

As in the case of the accelerating field, this term decreases inversely as the order n of the image.

15·7. Combined Electric and Magnetic Lenses. Any superposed axially symmetric electric and magnetic field constitutes a combined electric and magnetic electron lens. The equation of motion for paraxial electrons originally in a meridional plane (Eq. 15·61), together with the expression for the rotation of the image (Eq. 15·60) have been given. Just as in the case of the thin electric lens and the thin magnetic lens, the focal length of the thin combined lens can be obtained by integrating Eq. 15·62, utilizing the assumption that within the refracting field $R = r\Phi^{1/4}$ may be treated as a constant. Thus, by considering a paraxial electron incident parallel to the axis from object space,

$$\begin{aligned} \frac{1}{f_i} = -\frac{r_i'}{r_o} &= \left(\frac{\Phi_o}{\Phi_i}\right)^{1/4} \int_{z_o}^{z_i} T(z)\, dz \\ &= \left(\frac{\Phi_o}{\Phi_i}\right)^{1/4} \int_{z_o}^{z_i} \left[\frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 + \frac{0.022H^2}{\Phi}\right] dz \text{ cm}^{-1} \end{aligned} \qquad [15\cdot157]$$

Φ_o and Φ_i being the potentials in object and image space, respectively. The refractive power $1/f_i$ is, accordingly, simply the sum of the refractive powers of the component fields, with a suitably weighted value used for the potential in the calculation of the refractive power of the magnetic field.

The rotation of the image is again (Eq. 15·60):

$$\theta_i - \theta_o = 0.148 \int_{z_o}^{z_i} \frac{H}{\Phi^{1/2}}\, dz \qquad [15\cdot158]$$

and vanishes in the limit of a very thin lens.

The approximate method of path calculation resulting from dividing up the field into regions of constant field may also be adapted readily

to the use of path calculations in combined lenses (Fig. 15·26). In order to solve the equation

$$r'' = -\frac{\Phi'}{2\Phi}r' - \left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi}\right)r \qquad [15\cdot159]$$

it is convenient to introduce the new independent variable

$$\zeta = z - z_{ab} + \frac{\Phi_{ab}}{\Phi'} \qquad [15\cdot160]$$

where z_{ab} and Φ_{ab} are assumed to be the initial values of the z-coordinate and of the potential for the interval considered (that is, the interval b in Fig. 15·26).

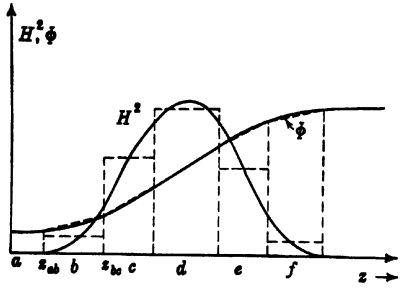


FIG. 15·26. Approximation-Field Representation for Calculating Electron Paths in Combined Lenses.

Equation 15·159 then becomes

$$r'' + \frac{r'}{2\zeta} + \frac{eH^2}{8mc^2\Phi'\zeta}r = 0 \qquad [15\cdot161]$$

It has the general solution:[18]

$$r = a\sin\left(\frac{eH^2\zeta}{2mc^2\Phi'}\right)^{1/2} + b\cos\left(\frac{eH^2\zeta}{2mc^2\Phi'}\right)^{1/2} \qquad [15\cdot162]$$

If the initial conditions $r = r_{ab}$, $r' = r'_{ab+}$ are substituted for $z = z_{ab}$ or $\zeta = \Phi_{ab}/\Phi'$, this becomes in terms of the original independent variable z:

$$r = r_{ab}\cos\left[\left(\frac{eH^2}{2mc^2\Phi'^2}\right)^{1/2}([(z - z_{ab})\Phi' + \Phi_{ab}]^{1/2} - \Phi_{ab}^{1/2})\right] \qquad [15\cdot163]$$
$$+ r'_{ab}\left(\frac{8mc^2\Phi_{ab}}{eH^2}\right)^{1/2}\sin\left[\left(\frac{eH^2}{2mc^2\Phi'^2}\right)^{1/2}([(z - z_{ab})\Phi' + \Phi_{ab}]^{1/2} - \Phi_{ab}^{1/2})\right]$$

[18] See Jahnke and Emde, reference 14, pp. 214 and 202.

The change in slope at the junctions between intervals is given, just as for a purely electric field, by

$$r'_{ab+} - r'_{ab-} = \frac{\Phi'_a - \Phi'_b}{4\Phi_{ab}} r_{ab} \qquad [15\cdot164]$$

and the rotation of the ray about the axis by

$$\theta - \theta_{ab} = \int_{\zeta_{ab}}^{\zeta} \left(\frac{eH^2}{8mc^2\Phi'_\zeta}\right)^{1/2} d\zeta$$

$$= \frac{H}{\Phi'}\left(\frac{e}{2mc^2}\right)^{1/2} ([(z - z_{ab})\Phi'_b + \Phi_{ab}]^{1/2} - \Phi_{ab}^{1/2}) \qquad [15\cdot165]$$

In general the study of combined electric and magnetic lenses requires recourse to the numerical integration of Eq. 15·159 or to the approximate procedure just outlined. Only in one case can the effects of the two fields be separated completely: that of superposed uniform electric and magnetic fields prevailing in the field intervals just considered in connection with the approximate method of path integration. The possibility of separating the effects of the two fields becomes evident only, however, if the time is introduced as parameter. Thus the motion in the direction of the z-axis of an electron leaving the origin at $t = 0$ is seen to be controlled entirely by the electric field, in accord with the formula for free fall:

$$z = \dot{z}_o t - \frac{eE}{2m} t^2 \qquad [15\cdot166]$$

whereas the projection of the motion on a plane normal to the axis is governed exclusively by the magnetic field:

$$r = \frac{2mc\dot{r}_o}{eH} \sin \frac{eHt}{2mc} \qquad [15\cdot167a]$$

$$\theta = \theta_o + \frac{eHt}{2mc} \qquad [15\cdot167b]$$

being a circle with a radius proportional to the initial velocity component normal to the axis $\dot{r}_o$ and inversely proportional to the strength of the magnetic field H. The path of the electron is thus a helix with progressively increasing pitch if E is an accelerating field. Erect images of unity magnification are formed by electrons with the same axial velocity

component at $eHt/(2mc) = \pi n$, that is, at

$$z_n = 2\pi n \frac{mc\dot{z}_o}{eH} - 2\pi^2 n^2 \frac{mc^2 E}{eH^2}$$

$$= 21.08n \left(\frac{\phi_{zo}}{H^2}\right)^{1/2} - 112.2n^2 \frac{E}{H^2} \text{ cm} \qquad [15\cdot168]$$

ϕ_{zo} being measured in volts, E in volts per centimeter, and H in gauss. ϕ_{zo} designates the accelerating voltage corresponding to the initial axial velocity of the electron.

If all the electrons enter the field with the same initial velocity, but at slightly varying inclinations $\tan \sigma_o = \dot{r}_o/\dot{z}_o = (\phi_r/\phi_{zo})^{1/2}$ (for $r_o = 0$), the electrons incident at the angle σ_o will focus ahead of those with $\sigma \doteq 0$ at the plane $z = z_n + \Delta z$:

$$\Delta z = \frac{dz}{d\phi_{zo}} \Delta\phi_{zo} = -\frac{dz}{d\phi_{zo}} \phi_r = -10.54n \frac{\phi_r}{\phi_{zo}} \left(\frac{\phi_{zo}}{H^2}\right)^{1/2} \qquad [15\cdot169]$$

forming a circle of diffusion of radius

$$\Delta r = \Delta z \frac{\dot{r}_o}{\dot{z}} = \Delta z \left(\frac{\phi_r}{\phi_{zo} - zE}\right)^{1/2} = -10.54n \frac{\phi_r^{3/2}}{H[\phi_{zo}(\phi_{zo} - zE)]^{1/2}} \qquad [15\cdot170]$$

in the plane $z = z_n$. ϕ_r is the accelerating voltage corresponding to the initial radial velocity $\dot{r}_o$. Equation 15·170 becomes identical with the corresponding expression for the uniform magnetic field (Eq. 15·33) alone if E is equal to zero.

Equations 15·169 and 15·170 assume that the angle σ_o is small, that is, that the initial axial velocity component is large compared to the initial tangential velocity component. An important case is that in which the initial tangential and axial velocities cover the same narrow range, as with the emission of electrons from a cathode, σ_o going up to $\pi/2$. In this case the axial deviation of focus for the electrons emitted tangentially becomes

$$\Delta z = -21.08n \left(\frac{\phi_r}{H^2}\right)^{1/2} \qquad [15\cdot171]$$

and the radius of the circle of diffusion

$$\Delta r = \Delta z \frac{\dot{r}_o}{\dot{z}} = \Delta z \left(\frac{\phi_r}{-z_n E}\right)^{1/2} = \frac{2\phi_r}{E} \qquad [15\cdot172]$$

Thus the circle of diffusion becomes independent of the order of focus.

REFERENCES

1. E. WIECHERT, "Velocity of cathode rays," *Wiedemanns Annalen*, Vol. 69, pp. 739–766, 1899.
2. H. BUSCH, "Calculation of the path of cathode rays in the axially symmetric electromagnetic field," *Ann. Physik*, Vol. 81, pp. 974–993, 1926.
3. H. BUSCH, "On the operation of the concentrating coil in the Braun tube," *Arch. Elektrotech.*, Vol. 18, pp. 583–594, June 1927.
4. W. GLASER, "Exact calculation of magnetic lenses with the field distribution $H = H_o/(1 + [z/a]^2)$," *Z. Physik*, Vol. 117, pp. 285–315, 1941.
5. I. G. MALOFF and D. W. EPSTEIN, *Electron Optics in Television*, McGraw-Hill Book Company, New York, 1938.
6. J. J. THOMSON, "Cathode rays," *Phil. Mag.*, Vol. 4, pp. 293–316., October 1897.
7. V. K. ZWORYKIN, G. A. MORTON, and L. MALTER, "The secondary-emission multiplier — a new electronic device," *Proc. Inst. Radio Engrs.*, Vol. 24, pp. 351–375, March 1936.
8. A. ROSE and H. IAMS, "Television pickup tubes using low-velocity electron beam," *Proc. Inst. Radio Engrs.*, Vol. 27, pp. 547–555, September 1939.
9. E. P. ADAMS, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, Smithsonian Institution, Washington, 1922.
10. A. ROSE, "A mechanical model for the motion of electrons in a magnetic field," *J. Applied Phys.*, Vol. 11, pp. 711–717, November 1940.
11. W. F. OSGOOD, *Mechanics*, The Macmillan Company, New York, 1937.
12. A. W. HULL, "The effect of a uniform magnetic field on the motion of electrons between coaxial cylinders," *Phys. Rev.*, Vol. 18, pp. 31–57, 1921.
13. A. ROSE, "Electron optics of cylindrical electric and magnetic fields," *Proc. Inst. Radio Engrs.*, Vol. 28, pp. 30–40, January 1940.
14. E. JAHNKE and F. EMDE, *Tables of Functions*, B. G. Teubner, Leipzig, 1933.

CHAPTER 16

ABERRATIONS OF ELECTRON LENSES

16·1. Classification of Aberrations. In Chapters 13 and 15 it has been demonstrated that, disregarding diffraction effects, any axially symmetric electric and magnetic field is capable of forming a true, geometrically faithful electron image of any object in a plane normal to the axis of symmetry of the field. However, the sharpness and faithfulness of the image so formed are contingent on the fulfilment of a number of conditions. They are: (1) The electron rays forming the image must make small angles with the axis or, in other words, the imaging aperture must be small. (2) The part of the object imaged must cover a small area about the axis or the image field must be small. (3) All the image-forming electrons leaving the object must have the same speed. If these conditions are not fulfilled the image will be, in general, neither sharp nor geometrically faithful. It is always possible, however, to construct a perfect, faithful image of a large object by assuming the same magnification, orientation, and sharpness for the image as a whole as for the small central portion about the axis imaged by narrow, *paraxial* pencils. This imaginary, perfect image is known as the *Gaussian* image of the object. The deviation of the actual image from this Gaussian image is termed the total aberration of the image. A specific vector value of the aberration may be ascribed to every imaging ray. It is the radius vector of the actual point of intersection of the ray with the Gaussian image plane from the Gaussian image point corresponding to the point of the object from which the ray considered originates (Fig. 16·1).

The total aberration of any ray may be regarded as the vector sum of a series of constituent aberrations which arise from different sources and are conveniently studied separately. They may be classified into geometrical aberrations, chromatic aberrations, space-charge defects, and the effects of disturbing fields and inaccuracies of construction and alignment.

The geometrical aberrations are the deviations of the ray intersections from the Gaussian image points which arise from the employment of imaging pencils of finite aperture and from the imaging of objects of finite extent in the absence of other disturbing factors. They are so named since they may be regarded as a function of the geometry of the refracting fields exclusively (for example, the shape of the equipotential surfaces).

The chromatic aberrations arise from differences in the initial velocities of the electrons. Just as variation in the color or wave length of light results in variations in the refractive power of glass lenses and other light-optical elements, so variations in initial velocity of the electrons correspond to differences in the refractive indices and the deflecting power of focusing fields.

A fixed amount of (axially symmetrically distributed) space charge of itself will influence the focusing properties of a refracting field, but, like a single fixed initial velocity, will not give rise to image defects. Varia-

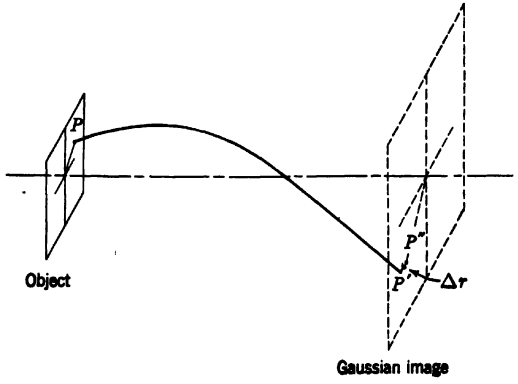


FIG. 16·1. The Aberration Δr for Any Ray PP' from the Gaussian Image Point P''.

tions in space charge or beam current, however, will cause such defects. Since it is rarely practical to keep the current perfectly constant and symmetrically distributed, the space-charge defect may in general be defined as the deviation of the ray intersection for the maximum space charge encountered from that for negligible space charge.

If the space-charge distribution is asymmetric it will act like the fourth source of aberration — external disturbing fields, mechanical misalignment, and constructional imperfections. All these destroy to some extent the basic condition for the lens property of electric and magnetic fields: axial symmetry. However, if the deviations from axial symmetry are small, they give rise to comparatively small deviations in the position of the ray intersections in the image plane.

16·2. The Derivation of the Geometrical Aberrations. The exact total geometrical aberration for any ray could, in principle, be found by calculating its exact path through the lens field either by means of the equations of motion (Eq. 15·6) or, more conveniently, with the aid of the path equations (15·54 and 15·56), and by determining the distance of its

intersection with the Gaussian image plane from the Gaussian image point of its point of origin in the object plane. The positions of the Gaussian image plane and image point would have to be determined separately, for example, by solving the paraxial ray equation,

$$r'' + \frac{r'\Phi'}{2\Phi} + r\left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi}\right) = 0 \qquad [16\cdot1]$$

$$\theta_i = \theta_o + \int_{z_o}^{z_i}\left(\frac{eH^2}{8mc^2\Phi}\right)^{\frac{1}{2}} dz = \theta_o + \chi(z_i) \qquad [16\cdot2]$$

for the two rays r_α and r_β with the initial conditions

$$r_\alpha(z_o) = 0 \qquad r'_\alpha(z_o) = 1 \qquad [16\cdot3a]$$

$$r_\beta(z_i) = 1 \qquad r'_\beta(z_i) = 0 \qquad [16\cdot3b]$$

with the image assumed to be in field-free space. The position of the image plane is fixed by the intersection of r_α with the axis: $r_\alpha(z_i) = 0$, whereas the magnification is given by $1/r_\beta(z_o)$ and the orientation of the image with respect to the object by $\chi(z_i) = \theta_i - \theta_o$. Thus, if the object point for any one ray is (r_P, θ_P, z_o), the corresponding Gaussian image point is $(r_P/r_\beta(z_o), \theta_P + \chi(z_i), z_i)$.

The procedure of finding the exact aberration for a single ray, outlined above, is, however, unsatisfactory from two points of view: (1) the exact calculation of the actual ray may meet great mathematical difficulties and require too detailed a knowledge of the electric and magnetic field distributions and (2) the result obtained is incapable of generalization. For these reasons it is customary to be content with an approximation to the actual aberration, based on the assumption that the apertures of the imaging pencils and the separations of the object points from the axis, although too large to permit disregarding all terms of order higher than the first in the lateral coordinates (the procedure leading to the paraxial equation and the Gaussian image), are yet small enough to render the added consideration of the third-order terms adequate. As was brought out in sections 13·1 and 15·3 the coordinates of the ray intersections in the image plane depend on odd-order terms in the object and aperture coordinates only. Accordingly, in what follows, only the *third-order aberrations* will be considered.

It may be remarked parenthetically that practical light optics, unlike electron optics, makes relatively little use of the third-order aberrations. The designer of light-optical instruments depends on the generally simple exact calculation of the actual rays. This becomes necessary since, because of the possibility of correcting the aberrations in light optics, aperture and object dimensions are frequently employed which render

the higher- (fifth-, seventh-, etc.) order aberration terms decidedly important. It will be seen below that similar possibilities of correction are lacking for the major electron-optical aberrations, resulting, in particular, in a restriction to small imaging apertures. This makes the third-order terms the predominant factor in determining the image quality under normal circumstances.

Formulas for the third-order aberrations have been derived in a number of different ways. Glaser[1] first adapted the methods familiar from light-optical theory to the problem, and Scherzer[2] obtained the same results directly from an integration of the path equation, retaining terms of the third order in the radial coordinate and the slope of the ray. The optical method, simplified in a manner suggested by Rogowski,[3] is outlined below.

The point of departure for this method is the Hamiltonian characteristic function, alternatively called *point iconal* or simply *ray function:*

$$W(x_o,y_o,z_o;\ x,y,z) = \int_{P_o(x_o,y_o,z_o)}^{P(x,y,z)} n\,ds \qquad [16\cdot4]$$

with

$$n = \phi^{1/2} - \left(\frac{e}{2mc^2}\right)^{1/2} A\cos\chi \qquad [16\cdot5]$$

where A is the value of the vector potential at the path element ds and χ the angle between it and the path element. The path of integration is such that

$$\delta W = 0 \qquad [16\cdot6]$$

that is, that the integral over the actual path is a minimum or assumes a stationary value as compared with neighboring paths of integration between P_o and P. W thus represents the *optical distance* between P_o and P, optical distance being defined as the geometrical distance along the actual path weighted by the index of refraction n.

If P_o is kept fixed, W = const describes a surface about P_o which may be regarded as the wave surface of the electron radiation diverging from P_o. For a purely electric field this is apparent, since here $n = \phi^{1/2} = \text{const}/\lambda$, where λ is the electron wave length, so that

$$W = \text{const}\int_{P_o}\frac{ds}{\lambda} = \text{const}' \qquad [16\cdot7]$$

denotes a surface of constant phase, the integral being equal to the num-

[1] See reference 1.
[2] See reference 2.
[3] See reference 3.

ber of wave lengths between P_o and the surface. In particular, for a region of constant index,

$$W = n\{(x - x_o)^2 + (y - y_o)^2 + (z - z_o)^2\}^{\frac{1}{2}} \qquad [16\cdot8]$$

so that here the wave surfaces are simply spheres about P_o. If the index is variable, that is, where the wave penetrates into a lens field, the sphere becomes distorted. An ideal lens would distort the wave surface into a spherical surface about another center, the image point (Fig. 16·2).

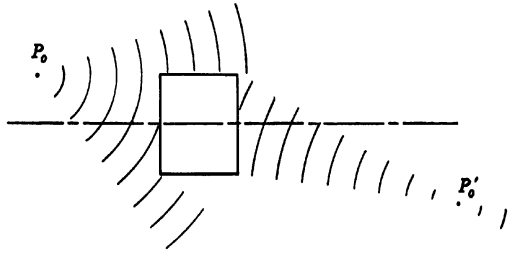


FIG. 16·2. Action of an Ideal Lens System on Wave Surfaces.

Actual lenses, on the other hand, will change the original spherical wave diverging from P_o into a wave which only approximates a sphere converging at the (Gaussian) image point, the deviation increasing with increasing distance of P_o from the axis and with increasing angular aperture of the imaging pencil (Fig. 16·3).

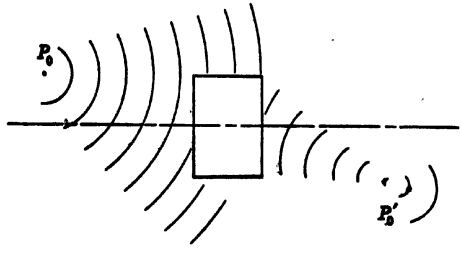


FIG. 16·3. Action of a Lens System with Aberrations on Wave Surfaces.

In isotropic regions (that is, regions free of magnetic field) the electron paths or electron rays from the point P_o are the orthogonal trajectories of the family of wave surfaces $W(x_o,y_o,z_o;\ x,y,z) = \text{const.}$ This follows from the fact that the rays between two points correspond to the minimum optical distance between those points. If, for example (Fig. 16·4), a certain ray passes through P_o and P_1 (on $W = W_1$) it will have to

intersect the adjoining equipotential $W = W_2$ in P_2, where P_1P_2 is normal to the equipotential surfaces, since P_1P_2 is smaller than any other possible path, such as P_1P_2'. (In the small region considered the index of refraction may be regarded as constant, so that optical and geometrical distances differ only by a constant factor.) Thus a normal to any wave surface in field-free image space is the continuation of some ray leaving the object point P_o, and the deviation of its intersection with the Gaussian image plane from the Gaussian image point P_i is the total geometric aberration of this ray. The property of the rays or electron paths of being wave-surface normals in image space will be utilized below to determine this aberration.

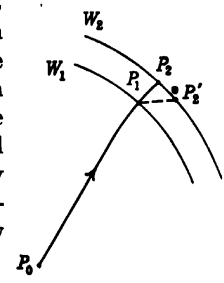


FIG. 16·4. The Rays as Normals to the Wave Surfaces.

Once again (see section 15·3, Fig. 15·4) consider an arbitrary axially symmetric electric and magnetic field, the object plane O (coordinates x_o,y_o,z_o), the image plane I (coordinates x_i,y_i,z_i), and near the image plane, in field-free space, the aperture plane A (coordinates x_a,y_a,z_a), z_o, z_i, and z_a being kept fixed. Furthermore, write

$$z_i - z_a = Z \qquad [16\cdot9]$$

The optical distance W between the points of the object plane and those of the aperture plane, being a function of the initial and final points only, can be written as a power series in the coordinates x_o,y_o,x_a,y_a:

$$W = W_0 + a_1x_o + a_2y_o + a_3x_a + a_4y_a + b_1x_o^2 + b_2x_oy_o + \cdots \qquad [16\cdot10]$$

$$= W_0 + W_1 + W_2 + W_3 + W_4 + \cdots$$

W_n denoting the part of the series containing terms of the nth order in the coordinates.

Now, in view of the axial symmetry, a rotation of the coordinates in both the object and the aperture plane through an identical angle θ does not change the optical distance W. W must remain invariant under the transformation

$$x_o \to x_o \cos\theta - y_o \sin\theta \qquad y_o \to x_o \sin\theta + y_o \cos\theta \qquad [16\cdot11a]$$

$$x_a \to x_a \cos\theta - y_a \sin\theta \qquad y_a \to x_a \sin\theta + y_a \cos\theta \qquad [16\cdot11b]$$

Since a rotation through 180 degrees changes the signs of all the coordinates, and hence changes the signs of the odd-order terms in W, these must vanish if W is to be unaffected by the transformation:

$$W_1 = W_3 = W_5 \cdots = 0 \qquad [16\cdot12]$$

If all coordinates are set equal to zero, $W = W_0$ becomes the optical distance from object to aperture plane along the axis:

$$W_0 = \int_{z_o}^{z_a} n(z)\,dz = \int_{z_o}^{z_a} \Phi^{1/2}(z)\,dz \qquad [16\cdot13]$$

$W_2, W_4, \cdots$ thus signify the deviation of the actual wave shape at the aperture plane from a plane wave normal to the axis. The first term may be written:

$$W_2 = b_1x_o^2 + b_2x_oy_o + b_3y_o^2 + b_4x_ox_a + b_5x_oy_a + b_6y_ox_a + b_7y_oy_a + b_8x_a^2 + b_9x_ay_a + b_{10}y_a^2 \qquad [16\cdot14]$$

Introducing the transformation (16·11) with $\theta = 90$ degrees leads to

$$W_2 = b_1y_o^2 - b_2x_oy_o + b_3x_o^2 + b_4y_oy_a - b_5y_ox_a - b_6x_oy_a + b_7x_ox_a + b_8y_a^2 - b_9x_ay_a + b_{10}x_a^2$$

or, with Eq. 16·14, $b_1 = b_3$; $b_2 = 0$; $b_4 = b_7$; $b_5 = -b_6$; $b_8 = b_{10}$; $b_9 = 0$, so that

$$W_2 = b_1(x_o^2 + y_o^2) + b_4(x_ox_a + y_oy_a) + b_5(x_oy_a - y_ox_a) + b_8(x_a^2 + y_a^2) \qquad [16\cdot15]$$

These four terms are the only second-order combinations of the four coordinates which are invariant under an arbitrary rotation. Hence all invariant fourth-order terms must be made up out of products of such terms, leading to the expression for W_4:

$$\begin{aligned} W_4 = {} & e_1(x_o^2 + y_o^2)^2 + e_2(x_o^2 + y_o^2)(x_ox_a + y_oy_a) \\ & + e_3(x_o^2 + y_o^2)(x_oy_a - y_ox_a) + e_4(x_o^2 + y_o^2)(x_a^2 + y_a^2) \\ & + e_5(x_ox_a + y_oy_a)^2 + e_6(x_ox_a + y_oy_a)(x_oy_a - y_ox_a). \\ & + e_7(x_oy_a - y_ox_a)^2 + e_8(x_ox_a + y_oy_a)(x_a^2 + y_a^2) \\ & + e_9(x_oy_a - y_ox_a)(x_a^2 + y_a^2) + e_{10}(x_a^2 + y_a^2)^2 \quad . \end{aligned} \qquad [16\cdot16]$$

The terms with e_4, e_5, and e_7 can be combined into two terms

$$\left(e_4 + \frac{e_5}{2} + \frac{e_7}{2}\right)(x_o^2 + y_o^2)(x_a^2 + y_a^2)$$

$$-\frac{1}{2}(e_5 - e_7)([y_ox_a - x_oy_a]^2 - [x_ox_a + y_oy_a]^2)$$

As in section 15·3, it is convenient to introduce a coordinate system in the object plane rotated through an angle χ_a with respect to the original

coordinate system in the object plane as well as to the coordinate systems in the aperture and image planes:

$$x_o' = \frac{b_4 x_o - b_5 y_o}{b_4'} \qquad y_o' = \frac{b_5 x_o + b_4 y_o}{b_4'}$$
$$b_4' = (b_4^2 + b_5^2)^{1/2} \qquad \tan \chi_a = -\frac{b_5}{b_4} \qquad [16\cdot17]$$

Then, with the new coordinates, it is possible to write

$$\begin{aligned} W = W_0 &+ b_1(x_o'^2 + y_o'^2) + b_4'(x_o' x_a + y_o' y_a) + b_8(x_a^2 + y_a^2) \\ &+ e_1'(x_o'^2 + y_o'^2)^2 + e_2'(x_o'^2 + y_o'^2)(x_o' x_a + y_o' y_a) \\ &+ e_3'(x_o'^2 + y_o'^2)(x_o' y_a - y_o' x_a) + e_4'(x_o'^2 + y_o'^2)(x_a^2 + y_a^2) \\ &+ e_5'([x_o' x_a + y_o' y_a]^2 - [x_o' y_a - y_o' x_a]^2) \\ &+ e_6'(x_o' y_a - y_o' x_a)(x_o' x_a + y_o' y_a) + e_7'(x_o' x_a + y_o' y_a)(x_a^2 + y_a^2) \\ &+ e_8'(x_o' y_a - y_o' x_a)(x_a^2 + y_a^2) + e_{10}(x_a^2 + y_a^2)^2 \\ &+ \text{terms of sixth and higher orders} \end{aligned} \qquad [16\cdot18]$$

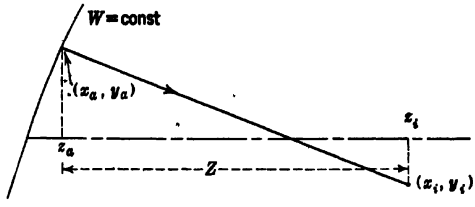


FIG. 16·5. Determination of Image-Plane Intersection from Wave Surface in Aperture Plane.

Consider now (Fig. 16·5) a particular point x_a, y_a in the aperture plane. The rays from the object point x_o', y_o' are the normals to the surface W = const through the point (x_a, y_a, z_a). The direction cosines of these normals to the surface are[4]

$$\frac{1}{n_i}\frac{\partial W}{\partial x_a} \qquad \frac{1}{n_i}\frac{\partial W}{\partial y_a} \qquad \frac{1}{n_i}\frac{\partial W}{\partial z_a}.$$

$n_i = \Phi_i^{1/2} = [(\partial W/\partial x_a)^2 + (\partial W/\partial y_a)^2 + (\partial W/\partial z_a)^2]^{1/2}$ being the gradient in image space of the ray function, in accordance with its definition in Eq. 16·4. Accordingly the location of the intersection with the image

[4] See Adams, reference 4, Formula 2.641.

plane of the actual ray through x_a,y_a is given by

$$x_i = x_a + \frac{Z}{n_i}\frac{\partial W}{\partial x_a}\left[1 + \frac{(x_i - x_a)^2}{Z^2} + \frac{(y_i - y_a)^2}{Z^2}\right]^{-\frac{1}{2}} \qquad [16\cdot19a]$$

$$y_i = y_a + \frac{Z}{n_i}\frac{\partial W}{\partial y_a}\left[1 + \frac{(x_i - x_a)^2}{Z^2} + \frac{(y_i - y_a)^2}{Z^2}\right]^{-\frac{1}{2}} \qquad [16\cdot19b]$$

The first-order terms in the expressions for x_i and y_i are

$$x_i^{(1)} = x_a + \frac{Z}{n_i} b_4' x_o' + \frac{2Z}{n_i} b_8 x_a \qquad [16\cdot20a]$$

$$y_i^{(1)} = y_a + \frac{Z}{n_i} b_4' y_o' + \frac{2Z}{n_i} b_8 y_a \qquad [16\cdot20b]$$

A comparison of these relations with Eqs. 15·44 shows that the coefficients of x_a,y_a must vanish, whereas those of x_o',y_o' must be equal to the magnification M:

$$b_8 = -\frac{n_i}{2Z} \qquad b_4' = \frac{Mn_i}{Z}$$

In other words, to this approximation a sharp, faithful image is obtained. The third-order terms yield the deviation of the actual ray intersections from the Gaussian image points:

$$\Delta x_i = x_i^{(3)} - x_i^{(1)} = \frac{Z}{n_i}\frac{\partial W_4}{\partial x_a} + \frac{1}{n_i}\left(\frac{(x_i^{(1)} - x_a)^2}{2Z} + \frac{(y_i^{(1)} - y_a)^2}{2Z}\right)\frac{\partial W_2}{\partial x_a} \qquad [16\cdot21a]$$

$$\Delta y_i = y_i^{(3)} - y_i^{(1)} = \frac{Z}{n_i}\frac{\partial W_4}{\partial y_a} + \frac{1}{n_i}\left(\frac{(x_i^{(1)} - x_a)^2}{2Z} + \frac{(y_i^{(1)} - y_a)^2}{2Z}\right)\frac{\partial W_2}{\partial y_a} \qquad [16\cdot21b]$$

where W_2 and W_4 are the second- and fourth-order terms, respectively, in W, Eq. 16·18. The second term on the right of Eqs. 16·21 is obtained by expanding the square root in Eqs. 16·19 and retaining the first terms of the expansion. If the fact that

$$x_i^{(1)} = Mx_o' \qquad y_i^{(1)} = My_o'$$

is utilized, the total aberration can be written

$$\begin{aligned}\Delta x_i = {} & S_1 x_o'(x_o'^2 + y_o'^2) - S_2 y_o'(x_o'^2 + y_o'^2) + S_3 x_a(x_o'^2 + y_o'^2) \\ & + S_4(x_a[x_o'^2 - y_o'^2] + 2x_o' y_o' y_a) \\ & + S_5(y_a[x_o'^2 - y_o'^2] - 2x_a x_o' y_o') + S_6(x_o'[3x_a^2 + y_a^2] + 2y_o' x_a y_a) \\ & + S_7(-y_o'[3x_a^2 + y_a^2] + 2x_o' x_a y_a) + S_8 x_a(x_a^2 + y_a^2)\end{aligned} \qquad [16\cdot22a]$$

$$\begin{aligned}\Delta y_i = {} & S_1 y_o'(x_o'^2 + y_o'^2) + S_2 x_o'(x_o'^2 + y_o'^2) + S_3 y_a(x'^2 + y_o'^2) \\ & + S_4(-y_a[x_o'^2 - y_o'^2] + 2x_a x_o' y_o') \\ & + S_5(x_a[x'^2 - y_o'^2] + 2y_a x_o' y_o') + S_6(y_o'[3y_a^2 + x_a^2] + 2x_o' x_a y_a) \\ & + S_7(x_o'[3y_a^2 + x_a^2] - 2y_o' x_a y_a) + S_8 y_a(x_a^2 + y_a^2)\end{aligned} \qquad [16\cdot22b]$$

with

$$S_1 = \frac{Ze_2'}{n_i} + \frac{M^3}{2Z^2} \qquad S_2 = \frac{Ze_3'}{n_i} \qquad S_3 = \frac{2Ze_4'}{n_i} - \frac{M^2}{Z^2}$$

$$S_4 = \frac{2Ze_5'}{n_i} - \frac{M^2}{2Z^2} \qquad S_5 = \frac{Ze_6'}{n_i} \qquad S_6 = \frac{Ze_7'}{n_i} + \frac{M}{2Z^2}$$

$$S_7 = \frac{Ze_8'}{n_i} \qquad S_8 = \frac{4Ze_{10}}{n_i} - \frac{1}{2Z^2}$$

It is thus seen that the deviations of the intersections of all rays with the image plane from the corresponding Gaussian image points may be determined from the eight coefficients $S_1 \cdots S_8$ and the coordinates of their intersections with the object plane and a suitably chosen aperture plane. The coefficients S_n, in turn, are functions of the field distribution in the electron-optical system and of the position of the object plane and the aperture plane. If the lens is purely electrostatic, the additional mirror symmetry about any plane through the axis causes S_2, S_5, and S_7 to vanish, leading to the same number and types of geometrical aberrations as occur in light optics. In this case, also, $\chi_a = 0$, $x_o' = x_o$, and $y_o' = y_o$, as has been seen in section 13·1. The component aberrations corresponding to the eight coefficients are described in detail below. In every case the aberration figure is derived for an object point on the y_o' axis ($x_o' = 0$). Furthermore, polar coordinates are introduced for the aperture plane:

$$x_a = r_a \cos\theta \qquad y_a = r_a \sin\theta \qquad [16\cdot23]$$

a. Distortion (Fig. 16·6). This aberration causes a shift of the image point proportional to the cube of the distance from the axis:

$$\Delta x = 0 \qquad \Delta y = S_1 y_o'^3 \qquad [16\cdot24]$$

It does not affect the sharpness, but only the faithfulness of the image. Depending on the sign of S_1, the distortion is either pincushion- or barrel-shaped, as shown in the figure.

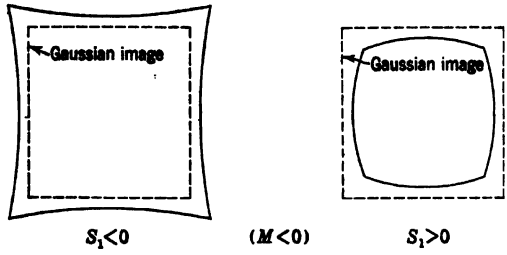


FIG. 16·6. Pincushion- and Barrel-Shaped Distortion.

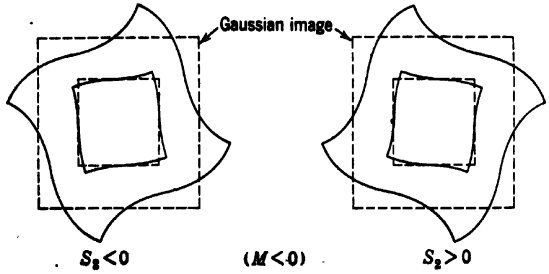


FIG. 16·7. Rotational Distortion.

b. Anisotropic Distortion (Fig. 16·7).

$$\Delta x = -S_2 y_o'^3 \qquad \Delta y = 0 \qquad [16\cdot25]$$

This occurs only in the presence of a magnetic field. The lateral shift, proportional to the cube of the separation of the object point from the axis, reverses sign with the direction of the magnetic field.

c. Curvature of Field (Fig. 16·8).

$$\Delta x = S_3 y_o'^2 r_a \cos\theta \qquad \Delta y = S_3 y_o'^2 r_a \sin\theta \qquad [16\cdot26]$$

The aberration figure is a circle of radius proportional to the square of the distance from the axis and to the first power of the radius of the aperture. The name is derived from the fact that, if it alone is present, a sharp image is formed on a curved surface tangent to the image plane on the axis. The radius of curvature R of the *image*

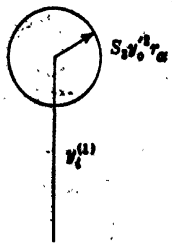


FIG. 16·8. Curvature of Field; Aberration Figure.

surface can be deduced from the geometry of Fig. 16·9:

$$IJ = \frac{Z}{r_a} \cdot S_3 y_o'^2 r_a = S_3 Z y_o'^2 \qquad R = \frac{M^2 y_o'^2}{2S_3 Z y_o'^2} = \frac{M^2}{2S_3 Z} \qquad [16\cdot27]$$

Depending on the sign of S_3, the image surface is concave or convex.

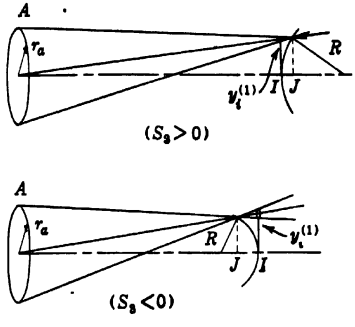


FIG. 16·9. Curvature of Field; Origin of Aberration Figure.

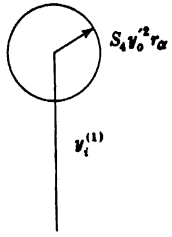


FIG. 16·10. Astigmatism, Aberration Figure.

d. Astigmatism (Fig. 16·10).

$$\Delta x = -S_4 y_o'^2 r_a \cos\theta \qquad \Delta y = S_4 y_o'^2 r_a \sin\theta \qquad [16\cdot28]$$

The aberration figure is again a circle with radius proportional to the square of the distance of the object point from the axis and to the first power of the aper ure radius. However, the sense of rotation around this circle is now opposite to that of the radius vector in the aperture plane. The formation of the aberration figure is shown in Fig. 16·11. The astigmatic pencil leaving the aperture plane converges more sharply in, for example, the meridional plane than in the tangential plane. Thus, at the two points of convergence, A and B, image lines rather than image points are formed, the two segments being mutually perpendicular. At a point midway between, in the image plane, the beam has a circular cross section of least diameter. The sets of line images A and B lie on two curved image surfaces — the *tangential* and *sagittal* image surfaces S_T and S_S, respectively. Except at the image plane and at the two image surfaces the beam cross section is elliptical. If curvature of field and astigmatism are simultaneously present, the aberration figure in the image plane becomes elliptical and the two image surfaces cease to be symmetrical with respect to the image plane.

e. Anisotropic Astigmatism (Fig. 16·12).

$$\Delta x = -S_5 y_o'^2 r_a \sin\theta \qquad \Delta y = -S_5 y_o'^2 r_a \cos\theta \qquad [16\cdot29]$$

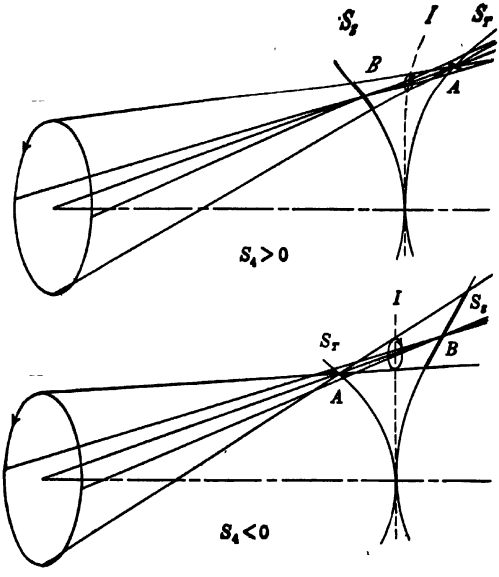


FIG. 16·11. Astigmatism; Origin of Aberration Figure.

This appears only in the presence of a magnetic field. The figure is again a circle with radius proportional to the square of the distance of the object point from the axis and to the first power of the aperture radius. It is traversed in a sense opposite to that of the ray intersection in the

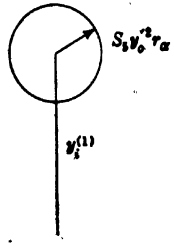


FIG. 16·12. Anisotropic Astigmatism; Aberration Figure.

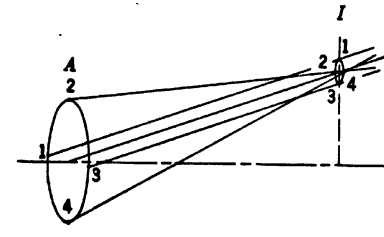


FIG. 16·13. Anisotropic Astigmatism; Origin of Aberration Figure.

aperture plane. In addition, there is a *phase difference* of 90 degrees, so that none of the rays of the pencil intersect anywhere, the least cross section of the beam occurring in the image plane (Fig. 16·13). In combination with ordinary astigmatism or curvature of field, the aberration figure becomes an ellipse of arbitrary orientation.

f. Coma (Fig. 16·14).

$$\Delta x = S_6 y_o' r_a^2 \sin 2\theta \qquad \Delta y = S_6 y_o' r_a^2 (2 - \cos 2\theta) \qquad [16\cdot30]$$

The aberration figure for fixed r_a is a circle of radius proportional to the square of the aperture radius and to the first power of the separation of

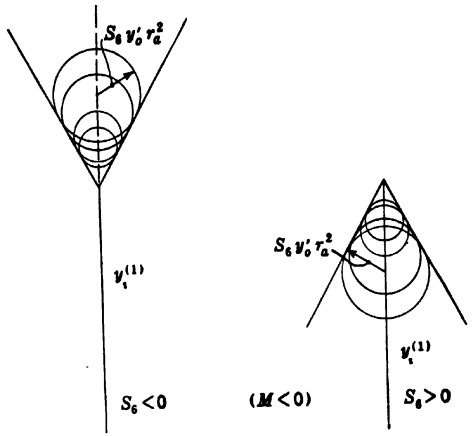


FIG. 16·14. Coma; Aberration Figure.

the object point from the axis, with its center a distance equal to its diameter from the Gaussian image point. The circles due to different zones of the aperture fill up a 60-degree angle. The resulting cometlike appearance of the image point is responsible for the name of the aberration, the vertex of the figure being the brightest portion thereof. Figure 16·15 shows the formation of the aberration figure. A single revolution about an aperture zone corresponds to two successive revolutions about one of the circles of the aberration figure.

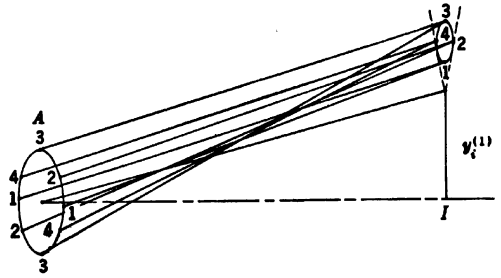


FIG. 16·15. Coma; Origin of Aberration Figure.

g. Anisotropic Coma (Fig. 16·16).

$$\Delta x = -S_7 y_o' r_a^2 (2 + \cos 2\theta) \qquad \Delta y = -S_7 y_o' r_a^2 \sin 2\theta \qquad [16\cdot31]$$

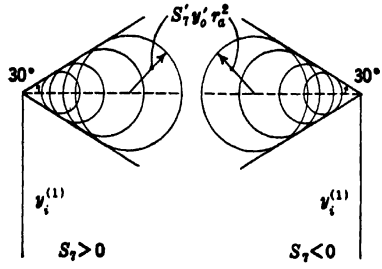


FIG. 16·16. Anisotropic Coma; Aberration Figure.

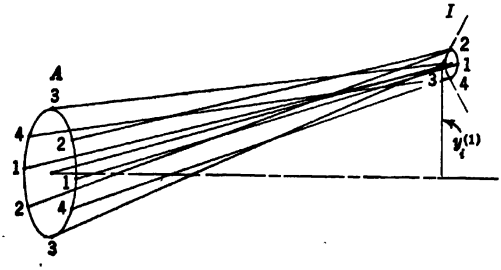


FIG. 16·17. Anisotropic Coma; Origin of Aberration Figure.

This aberration, present only with magnetic fields, leads to the figure of ordinary coma rotated through a right angle. Figure 16·17 shows the formation of the figure by the ray pencil leaving the aperture.



FIG. 16·18. Aperture Defect or Spherical Aberration; Aberration Figure.

h. Aperture Defect or Spherical Aberration (Fig. 16·18).

$$\Delta x = S_8 r_a^3 \cos\theta \qquad \Delta y = S_8 r_a^3 \sin\theta \qquad [16\cdot32]$$

The aperture defect is the same for all points of the object. The aberration figure is a circle about the Gaussian image point with a radius proportional to the cube of the aperture radius. Figure 16·19 shows the origin of the aberration figure, which results from the stronger convergence of the rays passing through the outer part of the aperture. The rays passing through the aperture zone r_a come to focus a distance $S_8 Z r_a^2$ from the image plane.

Two cautions with respect to the application of the above results to practical cases are not out of place:

1. The shape of the aberration figures depends (except for the case of distortion) on the electron distribution in the aperture plane. If, as is the case in electronic optics at times, the imaging pencil does not possess symmetry about the axis in the aperture plane chosen, the aberration figures will differ from those given above. They may of course still be determined by substituting the proper range of values r_a,θ in the aberration terms considered.

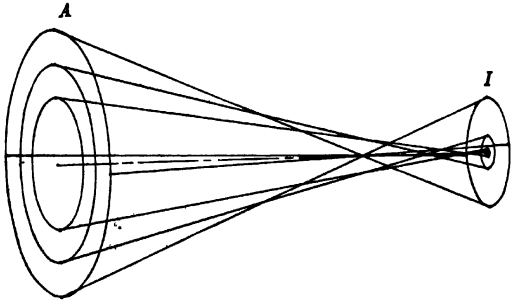


FIG. 16·19. Aperture Defect or Spherical Aberration; Origin of Aberration Figure.

2. The plane of best focus will not in general coincide exactly with the Gaussian image plane. Thus, by taking account of spherical aberration alone, the smallest circle of confusion is obtained in a plane a distance $d = 0.75S_8 r_{a\,\max}^2 \cdot Z$ from the Gaussian image plane. The circle of confusion is here only one-fourth as large as in the Gaussian image plane. Calculating the intensity distribution in different planes by the method of wave optics, Strehl[5] has shown that the plane of optimum focus lies somewhat closer to the Gaussian image plane, that is, at a distance $d = 0.5S_8 r_{a\,\max}^2 \cdot Z$. If the intensity distribution in the imaging pencil is not uniform, but has a maximum for $r_a = 0$, the focusing plane will lie yet closer to the Gaussian image plane.

16·3. Evaluation of the Geometrical Aberration Constants. It will be assumed that the paraxial ray equation (16·1) has been solved for the boundary conditions (Fig. 16·20):

$$r_\alpha(z_o) = 0 \qquad r'_\alpha(z_o) = 1 \qquad [16\cdot33a]$$

$$r_\gamma(z_o) = 1 \qquad r_\gamma(z_a) = 0 \qquad [16\cdot33b]$$

[5] See reference 5, p. 69.

$r_{a\alpha}$ will denote the distance of the ray r_α from the axis in the aperture plane. The rotation of the image will be given by

$$\chi_\alpha = \int_{z_o}^{z_a} \left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2} dz \qquad [16\cdot34]$$

since the region between the aperture and image planes is, in accordance with the definition of the former, field-free and so does not contribute to

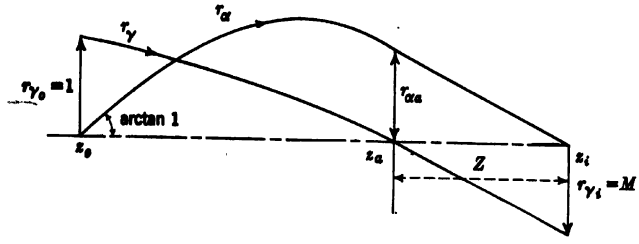


FIG. 16·20. The Solutions r_α and r_γ of the Paraxial Equation.

the integral. The path of an arbitrary ray leaving the object point r_o,θ_o and passing through the aperture plane in the point r_a,θ_a will be expressed by the complex function

$$w = x + iy = r\cdot e^{i\theta} \qquad [16\cdot35]$$

The location of the electron relative to a plane rotating, due to the presence of the magnetic field H, about the axis through the angle χ, is given by the function

$$u = w\cdot e^{-i\chi} = r\cdot e^{i(\theta-\chi)} \qquad [16\cdot36]$$

with

$$\chi = \int_{z_o}^{z} \left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2} dz$$

Utilizing Eqs. 15·58 and 15·59 for r'' and θ', u, for paraxial rays, is found to satisfy quite generally the equation

$$u'' + u'\frac{\Phi'}{2\Phi} + u\left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi}\right) = 0 \qquad [16\cdot37]$$

which is satisfied in particular by r_α and r_γ. Accordingly, u may be expressed as a linear combination of these two solutions:

$$u = u_o\cdot r_\gamma + \frac{u_a}{r_{a\alpha}}\cdot r_\alpha \qquad [16\cdot38]$$

u_o and u_a being the values of u in the object and in the aperture planes,

respectively. In terms of the values of the function $w = x + iy$ in these planes,

$$u = w_o \cdot r_\gamma + \frac{w_a}{r_{\alpha a}} \cdot e^{-i\chi_a} \cdot r_\alpha = \left\{ w_o' \cdot r_\gamma + \frac{w_a}{r_{\alpha a}} \cdot r_\alpha \right\} e^{-i\chi_a} \qquad [16\cdot39]$$

where w_o' represents the coordinates of the object point measured with respect to a system rotated through the angle χ_a relative to the original coordinate system.

Since the optical distance W between points in the object and aperture planes is to be expressed in terms of their respective coordinates w_o' and w_a, it is necessary to find the value of the angular momentum constant C (arising in the expression for θ') in terms of these quantities. Since, according to Eq. 15·59,

$$\theta' - \chi' = \frac{C}{r^2 \Phi^{1/2}} \qquad [16\cdot40]$$

it follows from Eq. 16·36 that

$$u'\bar{u} - \bar{u}'u = r\left(r' + \frac{iCr}{r^2\Phi^{1/2}}\right) - r\left(r' - \frac{iCr}{r^2\Phi^{1/2}}\right) = \frac{2iC}{\Phi^{1/2}} \qquad [16\cdot41]$$

On the other hand, by Eq. 16·39,

$$u'\bar{u} - \bar{u}'u = \frac{(w_o'\bar{w}_a - \bar{w}_o'w_a)(r_\gamma' r_\alpha - r_\alpha' r_\gamma)}{r_{\alpha a}} \qquad [16\cdot42]$$

Here $\bar{u}$, $\bar{w}_o'$, $\bar{w}_a$ denote, in conventional manner, the complex conjugates of u, w_o', and w_a, respectively. The value of the second factor on the right of Eq. 16·42 follows directly from the multiplication of the equation

$$r_\alpha'' + r_\alpha' \frac{\Phi'}{2\Phi} + r_\alpha \left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi} \right) = 0 \qquad [16\cdot43]$$

with r_γ, the multiplication of the corresponding equation for r_γ with r_α and their subtraction:

$$r_\gamma r_\alpha'' - r_\alpha r_\gamma'' = -(r_\gamma r_\alpha' - r_\alpha r_\gamma') \frac{\Phi'}{2\Phi} \qquad [16\cdot44]$$

which integrates to

$$\log (r_\gamma r_\alpha' - r_\alpha r_\gamma') = \log \frac{\text{const}}{\Phi^{1/2}} \qquad [16\cdot45]$$

Since the expression in parentheses on the left becomes unity for $z = z_o$,

$$r_\gamma r_\alpha' - r_\alpha r_\gamma' = \left(\frac{\Phi_o}{\Phi} \right)^{1/2} \qquad [16\cdot46]$$

so that

$$u'\bar{u} - \bar{u}'u = \left(\frac{\Phi_o}{\Phi}\right)^{\frac{1}{4}} \frac{\bar{w}_o' w_a - w_o' \bar{w}_a}{r_{aa}} \qquad [16\cdot47]$$

and

$$C = -i\Phi_o^{\frac{1}{4}} \frac{\bar{w}_o' w_a - w_o' \bar{w}_a}{2r_{aa}} \qquad [16\cdot48]$$

The expression in polar coordinates for the optical distance is

$$W = \int_{z_o}^{z_a} \left\{ \phi^{\frac{1}{4}}(1 + r^2\theta'^2 + r'^2)^{\frac{1}{4}} - \left(\frac{e}{2mc^2}\right)^{\frac{1}{2}} A_\theta r\theta' \right\} dz \qquad [16\cdot49]$$

This quantity is a minimum for the actual path between given points w_o' in the object and w_a in the aperture plane. The paraxial path solution, as given by Eq. 16·39, deviates from the actual path by quantities of the third order in w_o' and w_a, as has been shown in section 15·3. Since W is a minimum for the actual path, it increases, for infinitesimal deviations of the path, at most with the square of the deviation. Hence, if in place of the actual path the paraxial solution is substituted for r, r', and θ' in Eq. 16·49, the resulting error will be only of the sixth order in the coordinates w_o' and w_a. This will be without effect on the third-order deviations (corresponding to fourth-order terms in W) which are to be derived.

With the substitutions

$$\phi = \Phi - \frac{r^2}{4}\Phi'' + \frac{r^4}{64}\Phi^{\mathrm{IV}} = \Phi - \frac{u\bar{u}}{4}\Phi'' + \frac{(u\bar{u})^2}{64}\Phi^{\mathrm{IV}} \qquad [16\cdot50]$$

$$r^2\theta'^2 + r'^2 = w'\bar{w}' = u'\bar{u}' + \frac{eH^2}{8mc^2\Phi}u\bar{u} - i\left(\frac{eH^2}{8mc^2\Phi}\right)^{\frac{1}{4}}(u'\bar{u} - \bar{u}'u) \qquad [16\cdot51]$$

$$rA_\theta = \frac{r^2}{2}H - \frac{r^4}{16}H'' = \frac{u\bar{u}}{2}H - \frac{(u\bar{u})^2}{16}H'' \qquad [16\cdot52]$$

$$= \frac{C}{r^2\Phi^{\frac{1}{4}}} + \left(\frac{eH^2}{8mc^2\Phi}\right)^{\frac{1}{4}} = -i\frac{u'\bar{u} - \bar{u}'u}{2u\bar{u}} + \left(\frac{eH^2}{8mc^2\Phi}\right)^{\frac{1}{4}} \qquad [16\cdot53]$$

The expression for W may be separated into its terms W_0, W_2, W_4, etc., of the zero, second, fourth, etc., order in the coordinates:

$$W_0 = \int_{z_o}^{z_a} \Phi^{\frac{1}{4}}\, dz \qquad [16\cdot54]$$

$$
\begin{aligned}
W_2 &= \int_{z_o}^{z_a} \Phi^{1/2}\left\{-\frac{\Phi''}{8\Phi}u\bar{u} - \frac{eH^2}{16mc^2\Phi}u\bar{u} + \frac{u'\bar{u}'}{2}\right\}dz \\
&= \int_{z_o}^{z_a} \frac{d}{dz}\left(\frac{\Phi^{1/2}u'\bar{u}}{2}\right)dz = \left[\frac{\Phi^{1/2}u'\bar{u}}{2}\right]_{z_o}^{z_a} \\
&= \frac{1}{2}\left\{\Phi_i^{1/2}\left(w_o'r_{\gamma a}' + \frac{w_a r_{aa}'}{r_{aa}}\right)\bar{w}_a - \Phi_o^{1/2}\left(w_o'r_{\gamma o}' + \frac{w_a}{r_{aa}}\right)\bar{w}_o'\right\} \\
&= \frac{1}{2}\left\{-w_o'\bar{w}_o'r_{\gamma o}'\Phi_o^{1/2} - \frac{w_a\bar{w}_a\Phi_i^{1/2}}{Z} + (w_o'\bar{w}_a + \bar{w}_o'w_a)\Phi_i^{1/2}\frac{M}{Z}\right\}
\end{aligned}
\qquad [16\cdot55]
$$

The expressions for $r_{\gamma a}'$, r_{aa}, r_{aa}' follow from the geometry of Fig. 16·20. In addition, Eq. 16·46 is applied at $z = z_a$, yielding

$$
r_{aa}r_{\gamma a}' = -\left(\frac{\Phi_o}{\Phi_i}\right)^{1/2} \qquad [16\cdot56]
$$

Finally,

$$
\begin{aligned}
W_4 = \int_{z_o}^{z_a}\Big\{&-\frac{L}{4}(u\bar{u})^2 + \frac{iP\Phi^{1/2}}{2}u\bar{u}(u'\bar{u} - \bar{u}'u) - \frac{M'}{2}u\bar{u}u'\bar{u}' \\
&+ \frac{K\Phi}{4}(u'\bar{u} - \bar{u}'u)^2 + \frac{iQ\Phi^{1/2}}{2}u'\bar{u}'(u'\bar{u} - \bar{u}'u) - \frac{N}{4}(u'\bar{u}')^2\Big\}dz
\end{aligned}
\qquad [16\cdot57]
$$

where

$$
L = \frac{1}{32}\left\{-\frac{\Phi^{IV}}{\Phi^{1/2}} + \frac{\Phi''^2}{\Phi^{3/2}} + \frac{eH^2\Phi''}{mc^2\Phi^{3/2}} + \frac{e^2H^4}{4m^2c^4\Phi^{3/2}} - \frac{2eHH''}{mc^2\Phi^{3/2}}\right\} \qquad [16\cdot58a]
$$

$$
M' = \frac{1}{8}\left\{\frac{\Phi''}{\Phi^{1/2}} + \frac{eH^2}{2mc^2\Phi^{1/2}}\right\} \qquad [16\cdot58b]
$$

$$
K = \frac{eH^2}{16mc^2\Phi^{3/2}} \qquad [16\cdot58c]
$$

$$
Q = \left(\frac{e}{32mc^2}\right)^{1/2}\frac{H}{\Phi^{1/2}} \qquad [16\cdot58d]
$$

$$
N = \frac{\Phi^{1/2}}{2} \qquad [16\cdot58e]
$$

$$
P = \frac{1}{16}\left(\frac{e}{2m\Phi}\right)^{1/2}\left\{\frac{eH^3}{2mc^3\Phi} + \frac{\Phi''H}{c\Phi} - \frac{H''}{c}\right\} \qquad [16\cdot58f]
$$

are functions of the field distribution which were first introduced by Glaser.. By substituting the expression for u in terms of w_o' and w_a

given by Eq. 16·39, W_4 becomes:

$$\begin{aligned} W_4 = \frac{\Phi_i^{1/4}}{Z}\Big\{ & S_o'(w_o'\bar{w}_o')^2 + \frac{S_1'}{2} w_o'\bar{w}_o'(\bar{w}_o'w_a + w_o'\bar{w}_a) \\ & - i\frac{S_2'}{2} w_o'\bar{w}_o'(\bar{w}_o'w_a - w_o'\bar{w}_a) + \frac{S_3'}{2} w_o'\bar{w}_o'w_a\bar{w}_a \\ & + \frac{S_4'}{8}[(\bar{w}_o'w_a + w_o'\bar{w}_a)^2 + (\bar{w}_o'w_a - w_o'\bar{w}_a)^2] \\ & - i\frac{S_5'}{4}(\bar{w}_o'w_a + w_o'\bar{w}_a)(\bar{w}_o'w_a - w_o'\bar{w}_a) \\ & + \frac{S_6'}{2}(\bar{w}_o'w_a + w_o'\bar{w}_a)w_a\bar{w}_a \\ & - i\frac{S_7'}{2}(\bar{w}_o'w_a - w_o'\bar{w}_a)w_a\bar{w}_a + \frac{S_8'}{4}(w_a\bar{w}_a)^2\Big\} \end{aligned} \qquad [16\cdot59]$$

where

$$S_1' = \frac{M}{\Phi_o^{1/4}}\int_{z_o}^{z_a}\{r_\gamma^3 r_\alpha L + r_\gamma r_\gamma'(r_\alpha' r_\gamma + r_\alpha r_\gamma')M' + r_\gamma'^3 r_\alpha' N\}\,dz \qquad [16\cdot60a]$$

$$S_2' = M\int_{z_o}^{z_a}\{Pr_\gamma^2 + Qr_\gamma'^2\}\,dz \qquad [16\cdot60b]$$

$$S_3' = \frac{M}{r_{\alpha a}\Phi_o^{1/4}}\int_{z_o}^{z_a}\{2Lr_\gamma^2 r_\alpha^2 + M'(r_\alpha r_\gamma' + r_\gamma r_\alpha')^2 + 2Nr_\gamma'^2 r_\alpha'^2 + \Phi_o K\}\,dz \qquad [16\cdot60c]$$

$$S_4' = \frac{M}{r_{\alpha a}\Phi_o^{1/4}}\int_{z_o}^{z_a}\{Lr_\gamma^2 r_\alpha^2 + 2M' r_\gamma r_\alpha r_\gamma' r_\alpha' + Nr_\gamma'^2 r_\alpha'^2 - \Phi_o K\}\,dz \qquad [16\cdot60d]$$

$$S_5' = \frac{2M}{r_{\alpha a}}\int_{z_o}^{z_a}\{Pr_\gamma r_\alpha + Qr_\gamma' r_\alpha'\}\,dz \qquad [16\cdot60e]$$

$$S_6' = \frac{M}{r_{\alpha a}^2\Phi_o^{1/4}}\int_{z_o}^{z_a}\{Lr_\gamma r_\alpha^3 + M' r_\alpha r_\alpha'(r_\alpha' r_\gamma + r_\gamma' r_\alpha) + Nr_\gamma' r_\alpha'^3\}\,dz \qquad [16\cdot60f]$$

$$S_7' = \frac{M}{r_{\alpha a}^2}\int_{z_o}^{z_a}\{Pr_\alpha^2 + Qr_\alpha'^2\}\,dz \qquad [16\cdot60g]$$

$$S_8' = \frac{M}{r_{\alpha a}^2\Phi_o^{1/4}}\int_{z_o}^{z_a}\{Lr_\alpha^4 + 2M' r_\alpha^2 r_\alpha'^2 + Nr_\alpha'^4\}\,dz \qquad [16\cdot60h]$$

The deviation of the ray intersection from the Gaussian image point is

obtained by the differentiation of Eqs. 16·59 and 16·55, as indicated by Eq. 16·21. Hence:

$$\Delta w_i = \Delta x + i\Delta y = \left(S_1' + iS_2' + \frac{M^3}{2Z^2}\right) w_o'^2 \bar{w}_o' + \left(S_3' - \frac{M^2}{Z^2}\right) w_o' \bar{w}_o' w_a + \left(S_4' + iS_5' - \frac{M^2}{2Z^2}\right) w_o'^2 \bar{w}_a + \left(S_6' - iS_7' + \frac{M}{2Z^2}\right) \bar{w}_o' w_a^2 + 2\left(S_6' + iS_7' + \frac{M}{2Z^2}\right) w_o' w_a \bar{w}_a + \left(S_8' - \frac{1}{2Z^2}\right) \bar{w}_a w_a^2 \qquad [16·61]$$

The only terms under the integrals which do not vanish in the field-free region between z_a and z_i (aperture and image planes) are those with the coefficient N, Eq. 16·58e. In this region, furthermore, all these terms are constant, with

$$r_\alpha' = -\frac{r_{\alpha a}}{Z} \qquad r_\gamma' = \frac{M}{Z}$$

$$r_{\alpha a} = -\frac{1}{r_{\gamma a}'}\left(\frac{\Phi_o}{\Phi_i}\right)^{1/2} = -\frac{Z}{M}\left(\frac{\Phi_o}{\Phi_i}\right)^{1/2} \qquad [16·62]$$

Thus, for example, for S_1',

$$\frac{M}{\Phi_o^{1/2}} \int_{z_a}^{z_i} r_\gamma'^3 r_\alpha' N \, dz = \frac{M}{\Phi_o^{1/2}} \frac{M^3}{Z^3} \frac{-r_{\alpha a}}{Z} \frac{\Phi_i^{1/2}}{2} Z = \frac{M^3}{2Z^2}$$

Accordingly, $S_1' + M^3/(2Z^2)$ is equal to the expression resulting from S_1' (Eq. 16·60a) if the upper limit of the integral is made z_i in place of z_o. The same will be found to hold for the remaining aberration coefficients. Thus

$$S_1 = S_1' + \frac{M^3}{2Z^2} \qquad S_2 = S_2' \qquad S_3 = S_3' - \frac{M^2}{Z^2}$$

$$S_4 = S_4' - \frac{M^2}{2Z^2} \qquad S_5 = S_5' \qquad S_6 = S_6' + \frac{M}{2Z^2} \qquad [16·63]$$

$$S_7 = S_7' \qquad S_8 = S_8' - \frac{1}{2Z^2}$$

are obtained simply by replacing the upper limit z_a in the integrals in Eqs. 16·60 by z_i.

These equations for $S_1 \cdots S_8$ are not particularly convenient for calculation, since they involve the fourth derivative of the electric potential and the second derivative of the magnetic field along the axis.

These derivatives are difficult to obtain in practice, especially if, as is usually the case, the fields are measured experimentally and the differentiation has to be carried out by graphical or numerical methods. By repeated integrations by parts it is possible, however, by utilizing the fact that r_α vanishes for both $z = z_o$ and $z = z_i$, to throw the equations for the aberration coefficients into a form in which only the first and second derivatives of the potential and the first derivative of the magnetic field appear, apart from the potential and the magnetic field themselves. Since these formulas are extremely useful, they are given below, in the form obtained by Scherzer:[6]

$$\overline{S_1} = \frac{M}{16\Phi_o^{1/2}}\int_{z_o}^{z_i}\Phi^{-3/4}r_\gamma^3 r_\alpha\left\{U + V\left(3\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha}\right) + W\left(\frac{r'^2_\gamma}{r^2_\gamma} + \frac{r'_\gamma r'_\alpha}{r_\gamma r_\alpha}\right)\right\}dz$$
$$+ \frac{M}{4}\left[r_\gamma^2\left(\frac{\Phi''}{8\Phi} + \frac{5}{32}\frac{\Phi'^2}{\Phi^2} + \frac{\Phi' r'_\gamma}{2\Phi r_\gamma} + \frac{3r'^2_\gamma}{2r^2_\gamma} + \frac{eH^2}{8mc^2\Phi}\right)\right]_{z_o}^{z_i} \qquad [16\cdot64a]$$

$$S_2 = \frac{M}{16}\left(\frac{2e}{mc^2}\right)^{1/2}\int_{z_o}^{z_i}\Phi^{-3/4}Hr_\gamma^2\left\{X - \frac{\Phi'}{2}\frac{r'_\gamma}{r_\gamma} + \Phi\frac{r'^2_\gamma}{r^2_\gamma}\right\}dz$$
$$+ \frac{M}{32}\left(\frac{2e}{mc^2}\right)^{1/2}\left[\Phi^{-1/4}Hr_\gamma^2\left(\frac{3\Phi'}{2\Phi} + \frac{2r'_\gamma}{r_\gamma} - \frac{H'}{H}\right)\right]_{z_o}^{z_i} \qquad [16\cdot64b]$$

$$S_3 = \frac{M}{8r_{\alpha a}\Phi_o^{1/2}}\int_{z_o}^{z_i}\Phi^{-3/4}r_\gamma^2 r_\alpha\left\{U + 2V\left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha}\right)\right.$$
$$\left. - \left(W + \frac{\Phi'^2}{2}\right)\left(\frac{r'_\gamma}{r_\gamma} - \frac{r'_\alpha}{r_\alpha}\right)^2 + 2W\frac{r'_\gamma r'_\alpha}{r_\gamma r_\alpha}\right\}dz \qquad [16\cdot64c]$$
$$+ \frac{M}{8r_{\alpha a}}\left[\Phi_o^{1/4}\Phi^{-3/4}\Phi' + 4r'_\gamma r'_\alpha\right]_{z_o}^{z_i}$$

$$S_4 = \frac{M}{16r_{\alpha a}\Phi_o^{1/2}}\int_{z_o}^{z_i}\Phi^{-3/4}r_\gamma^2 r_\alpha^2\left\{U + 2V\left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha}\right)\right.$$
$$\left. + (3W + \Phi'^2)\left(\frac{r'_\gamma}{r_\gamma} - \frac{r'_\alpha}{r_\alpha}\right)^2 + 2W\frac{r'_\gamma r'_\alpha}{r_\gamma r_\alpha}\right\}dz + \frac{M}{4r_{\alpha a}}\left[r'_\gamma r'_\alpha\right]_{z_o}^{z_i} \qquad [16\cdot64d]$$

$$S_5 = \frac{M}{8r_{\alpha a}}\left(\frac{2e}{mc^2}\right)^{1/2}\int_{z_o}^{z_i}\Phi^{-1/4}Hr_\gamma r_\alpha\left\{X - \frac{\Phi'}{4}\left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha}\right)\right.$$
$$\left. + \Phi\frac{r'_\gamma r'_\alpha}{r_\gamma r_\alpha}\right\}dz + \frac{M}{16r_{\alpha a}}\left(\frac{2e\Phi_o}{mc^2}\right)^{1/2}\left[\frac{H}{\Phi}\right]_{z_o}^{z_i} \qquad [16\cdot64e]$$

[6] See reference 6.

$$S_6 = \frac{M}{16r_{\alpha a}^2\Phi_o^{1/2}}\int_{z_o}^{z_i}\Phi^{-3/2}r_\gamma r_\alpha^3\left\{U + V\left(\frac{r'_\gamma}{r_\gamma} + \frac{3r'_\alpha}{r_\alpha}\right) + W\left(\frac{r'_\gamma r'_\alpha}{r_\gamma r_\alpha} + \frac{r'^2_\alpha}{r_\alpha^2}\right)\right\}dz + \frac{M}{8r_{\alpha a}^2}\left[r'^2_\alpha\right]_{z_o}^{z_i} \qquad [16\cdot64f]$$

$$S_7 = \frac{M}{16r_{\alpha a}^2}\left(\frac{2e}{mc^2}\right)^{1/2}\int_{z_o}^{z_i}\Phi^{-3/2}Hr_\alpha^2\left\{X - \frac{\Phi'}{2}\frac{r'_\alpha}{r_\alpha} + \Phi\frac{r'^2_\alpha}{r_\alpha^2}\right\}dz \qquad [16\cdot64g]$$

$$S_8 = \frac{M}{16r_{\alpha a}^3\Phi_o^{1/2}}\int_{z_o}^{z_i}\Phi^{-3/2}r_\alpha^4\left\{U + 4V\frac{r'_\alpha}{r_\alpha} + 2W\frac{r'^2_\alpha}{r_\alpha^2}\right\}dz \qquad [16\cdot64h]$$

with

$$U = \frac{5\Phi''^2}{4} + \frac{5\Phi'^4}{24\Phi^2} + \frac{e\Phi H'^2}{mc^2} + \frac{3e^2H^4}{8m^2c^4} + \frac{35e\Phi'^2H^2}{16mc^2\Phi} - \frac{3e\Phi'HH'}{mc^2} \qquad [16\cdot64i]$$

$$V = \frac{7\Phi'^3}{6\Phi} - \frac{e\Phi'H^2}{2mc^2} \qquad [16\cdot64j]$$

$$W = -\frac{3\Phi'^2}{4} - \frac{e\Phi H^2}{2mc^2} \qquad [16\cdot64k]$$

$$X = \frac{3eH^2}{8mc^2} + \frac{9\Phi'^2}{8\Phi} - \frac{\Phi'H'}{H} \qquad [16\cdot64l]$$

16·4. Validity of the Aberration Expressions. From the derivation of the aberration expressions, with a field-free space postulated between aperture and image plane, it might be inferred that the formulas obtained are valid only when the image is in field-free space. This is not so. If the image plane is located in a region which is not field-free, the field may be regarded as cut off sharply at the image plane, the region beyond being field-free and at the potential $\Phi_i = \Phi(z_i)$. The "aperture plane" is located in this region beyond the image. Under these circumstances, the derivation of the aberration expressions remains valid. The imagined modification of the field beyond the image plane cannot affect the result, since the electron paths between object and image are not influenced thereby.

The formulas (16·60) are applicable to all continuous fields without space charge, producing either real images or virtual images. The Eqs. 16·64, on the other hand, apply only to fields producing real images and those fields producing virtual images for which the aperture plane may be located in field-free space without cutting off the field

discontinuously at some point. The vanishing of either $r_a(z_i)$ or of H and Φ' and their derivatives at $z = z_a$ is required to obtain the second set of formulas.

An example of the second class of fields to which Eqs. 16·64 apply is an electron microscope projector lens of short focus, for which the object is virtual, the image real. The interchange of object and image is immaterial, since the aberrations resulting for the two arrangements differ only by a constant factor equal to the magnification.

In every case the integrations in Eqs. 16·64 must be carried out over the entire imaging field. In Eqs. 16·60 the integration is first carried out through the entire imaging field and then, from the end of the imaging field, to the image point $z = z_i$. Only the term in the integrand with the coefficient N contributes to the last part of the integration. Neither Eqs. 16·60 nor Eqs. 16·64 are, of course, applicable to the determination of the aberrations of electron mirrors or cathode lenses, since here, for some point of the field, $\Phi = 0$, making terms in the integrands of all the formulas infinite.

16·5. Geometric Aberrations of the Electron Mirror. In order to determine the aberrations of the electron mirror it is necessary, to avoid singularities such as the infinity of r' at the reversal point, to obtain a parametric representation of the electron paths. The natural parameter to use for this purpose is the time. The general equations of motion in an axially symmetric field:

$$\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta}) = \frac{e}{mc}(\dot{r}h_z - \dot{z}h_r) \qquad [16\cdot65a]$$

$$\ddot{r} - r\dot{\theta}^2 = \frac{e}{m}\frac{\partial\phi}{\partial r} - \frac{e}{mc}r\dot{\theta}h_z \qquad [16\cdot65b]$$

$$\ddot{z} = \frac{e}{m}\frac{\partial\phi}{\partial z} + \frac{e}{mc}r\dot{\theta}h_r \qquad [16\cdot65c]$$

where a dot, in conventional manner, indicates differentiation with respect to the time, yield, to a first approximation, since by Eq. 15·39

$$\dot{\theta} = \frac{C'}{r^2} + \frac{eH}{2mc} - \frac{eH''}{16mc}r^2 + \cdots \qquad [16\cdot66]$$

and, according to Eqs. 14·41 and 14·42,

$$h_z = H - \frac{r^2H''}{4}\cdots \qquad [16\cdot67a]$$

$$h_r = -\frac{rH'}{2} + \frac{r^3H'''}{16}\cdots \qquad [16\cdot67b]$$

the differential equations

$$\ddot{r} = \frac{C'^2}{r^3} - \left(\frac{e^2H^2}{4m^2c^2} + \frac{e\Phi''}{2m}\right)r \qquad [16\cdot68]$$

$$\dot{z} = \left(\frac{2e\Phi}{m}\right)^{1/2} \qquad [16\cdot69]$$

As before, it is convenient to introduce as dependent variable the complex quantity

$$u = r\cdot\exp\left[i\left(\theta - \int_{t_o}^{t}\frac{eH}{2mc}dt\right)\right] \qquad [16\cdot70]$$

Then, also to a first approximation,

$$\ddot{u} = -\left(\frac{e^2H^2}{4m^2c^2} + \frac{e\Phi''}{2m}\right)u \qquad [16\cdot71]$$

If r_α and r_γ are the two particular solutions of Eq. 16·71 with the boundary conditions

$$r_\alpha(z_o) = 0 \qquad \dot{r}_\alpha(z_o) = \left(\frac{2e\Phi_o}{m}\right)^{1/2} \qquad [16\cdot72a]$$

$$r_\gamma(z_o) = 1 \qquad r_\gamma(z_a) = 0 \qquad [16\cdot72b]$$

where z_a is the abscissa of the " aperture plane," defined as before as a plane normal to the axis with the region between it and the image plane field-free, then any solution of Eq. 16·71 may be written

$$u = u_o r_\gamma + \frac{u_a r_\alpha}{r_{\alpha a}} = \left(w'_o r_\gamma + \frac{w_a r_\alpha}{r_{\alpha a}}\right)e^{-i\chi_a} \qquad [16\cdot73]$$

where

$$\chi_a = \int_{t_o}^{t_a}\frac{eH}{2mc}dt$$

and $w'_o = x'_o + iy'_o$ represents the coordinates in the object plane relative to a system rotated through the angle χ_a with respect to the coordinate systems in the aperture and image planes. From Eq. 16·70 it follows that

$$\dot{u}\bar{u} - \dot{\bar{u}}u = 2iC'$$

If the quantities to the third order in the coordinates are retained, the

differential equations become:

$$\ddot{u}_s = -\left(\frac{e\Phi_s''}{2m} + \frac{e^2H_s^2}{4m^2c^2}\right)u_s + \left(\frac{e\Phi^{IV}}{16m} + \frac{e^2HH''}{8m^2c^2}\right)u^2\bar{u}$$
$$-\frac{ie}{mc}\left\{\frac{H''\dot{u}u\bar{u}}{4} + \frac{H'''\dot{z}u^2\bar{u}}{16} - H'\Delta z\dot{u} - \frac{H'\Delta\dot{z}u}{2}\right. \qquad [16\cdot74]$$
$$\left. -\frac{H''\dot{z}\Delta zu}{2} + \frac{H'\Delta z(\dot{u}\bar{u} - \dot{\bar{u}}u)}{2\bar{u}}\right\}$$

and, for $\Phi \neq 0$,

$$\dot{z}_s = \left(\frac{2e\phi_s}{m} - \dot{r}^2 - r^2\dot{\theta}^2\right)^{1/2} = \left(\frac{2e\Phi_s}{m}\right)^{1/2}\left(1 - \frac{\Phi''w\bar{w}}{8\Phi} - \frac{m\dot{w}\dot{\bar{w}}}{4e\Phi}\right) \qquad [16\cdot75]$$
$$= \left(\frac{2e\Phi_s}{m}\right)^{1/2}\left(1 - \frac{\Phi''u\bar{u}}{8\Phi} - \frac{eH^2u\bar{u}}{16mc^2\Phi} - \frac{m\dot{u}\dot{\bar{u}}}{4e\Phi} + \frac{iH(\dot{u}\bar{u} - u\dot{\bar{u}})}{8\Phi}\right)$$

In these equations the subscript s denotes the second approximation, whereas no subscript is given to first-approximation quantities: for example, $\Phi_s'' = \Phi''(z_s)$, whereas $\Phi'' = \Phi''(z)$. The differences between the first and second approximation

$$\Delta u = u_s - u \qquad \Delta z = z_s - z \qquad [16\cdot76]$$

yield the aberration

$$\Delta u_i = \Delta u(t_i) - \Delta z(t_i)\cdot\dot{u}(t_i)\cdot\left(\frac{2e\Phi_i}{m}\right)^{-1/2} \qquad [16\cdot77]$$

t_i is such that $r_a(t_i) = 0$; $t_i - t_o$ represents the time required for a paraxial electron to reach the image plane. The equation for Δu becomes, from Eq. 16·74:

$$\Delta\ddot{u} + \left(\frac{e\Phi''}{2m} + \frac{e^2H^2}{4m^2c^2}\right)\Delta u = R(t) \qquad [16\cdot78]$$

$$R(t) = -\left(\frac{e\Phi''}{2m} + \frac{e^2HH'}{2m^2c^2}\right)u\Delta z + \left(\frac{e\Phi^{IV}}{16m} + \frac{e^2HH''}{8m^2c^2}\right)u^2\bar{u}$$
$$-\frac{ie}{mc}\left\{\frac{H''\dot{u}u\bar{u}}{4} + \frac{H'''\dot{z}u^2\bar{u}}{16} - H'\Delta z\dot{u} - \frac{H'\Delta\dot{z}u}{2}\right. \qquad [16\cdot79]$$
$$\left. -\frac{H''\dot{z}\Delta zu}{2} + \frac{H'\Delta z(\dot{u}\bar{u} - \dot{\bar{u}}u)}{2\bar{u}}\right\}$$

This differential equation may be integrated[7] to yield

$$\Delta u(t_i) = -\left(\frac{2e\Phi_o}{m}\right)^{-\frac{1}{2}} M \int_{t_o}^{t_i} R(\tau) r_a(\tau)\, d\tau \qquad [16\cdot80]$$

The equation for Δz becomes, on the other hand,

$$\Delta\dot{z} = \left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}} \left(\frac{\Phi'\Delta z}{2\Phi} - Q(t)\right)$$

or

$$\frac{d}{dt}\left[\frac{\Delta z}{\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}}\right] = -Q(t) \qquad [16\cdot81]$$

with

$$Q(t) = \left(\frac{\Phi''}{8\Phi} + \frac{eH^2}{16mc^2\Phi}\right) u\bar{u} + \frac{m\dot{u}\dot{\bar{u}}}{4e\Phi} - \frac{iH(\dot{u}\bar{u} - u\dot{\bar{u}})}{8c\Phi}$$

Hence, for $t < t_r$, t_r being the instant of reversal for the paraxial electrons leaving the object at the time t_o,

$$\Delta z = -\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}} \int_{t_o}^{t} Q(t)\, dt \qquad [16\cdot82]$$

On the other hand, at $t = t_r$ the value of Δz is such that the total kinetic energy is associated with the velocity of the electron in a plane normal to the axis:

$$e\phi_r - e\Phi_r'\Delta z = \frac{m\dot{w}_r\dot{\bar{w}}_r}{2} \qquad [16\cdot83]$$

Hence,

$$\Delta z_r = \left[\frac{2\Phi}{\Phi'} Q(t)\right]_{t=t_r} \qquad [16\cdot84]$$

Since the integrand in Eq. 16·82 becomes infinite at $t = t_r$ ($\Phi_r = 0$), it is convenient to transform it by an integration by parts. Thus the expression

$$\begin{aligned}\Delta z &= \left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}} \left\{\int_{t_r}^{t} \frac{2\Phi}{\Phi'} Q(t) \frac{d}{dt}\left(\frac{m}{2e\Phi}\right)^{\frac{1}{2}} dt + \lim_{t\to t_r} \frac{\Delta z_r}{\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}}\right\} \\ &= \frac{2\Phi}{\Phi'} Q(t) - \left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}} \int_{t_r}^{t} \left\{\left(\frac{\Phi'''}{4\Phi'} - \frac{\Phi'^2}{4\Phi'^2} + \frac{eHH'}{4mc^2\Phi'} - \frac{eH^2\Phi''}{8mc^2\Phi'^2}\right) u\bar{u}\right. \\ &\qquad \left. - \frac{m\Phi''}{2e\Phi'^2} \dot{u}\dot{\bar{u}} - i\left(\frac{H'}{4\Phi'} - \frac{H\Phi''}{4\Phi'^2}\right) \frac{\dot{u}\bar{u} - u\dot{\bar{u}}}{c}\right\} dt\end{aligned} \qquad [16\cdot85]$$

[7] See Adams, reference 4, Formula 8.410.

contains only finite terms near $t = t_r$, provided that $\Phi_r' \neq 0$, and satisfies the differential equation.

Equation 16·85 would logically be used for a region between t_1 and t_2, $t_1 < t_r < t_2$, where $\Phi' < 0$ and Φ is small. Beyond t_2, similarly as for the region $t < t_1$,

$$\Delta z = \left(\frac{\Phi}{\Phi_2}\right)^{\frac{1}{4}} \Delta z_2 - \left(\frac{2e\Phi}{m}\right)^{\frac{1}{4}} \int_{t_2}^{t} Q(t)\, dt \qquad [16\cdot86]$$

Substitution in the equations for $\Delta u(t_i)$ and Δu_i (Eqs. 16·80 and 16·77) of the expressions for Δz in Eqs. 16·82, 16·85, and 16·86 leads to an expression of the third-order aberration of the electron mirror in terms of u and $\dot{u}$. If here

$$u = e^{-i\chi_a}\left(w_o' r_\gamma + \frac{w_a r_\alpha}{r_{\alpha a}}\right) \qquad [16\cdot87a]$$

$$\dot{u} = e^{-i\chi_a}\left(w_o' \dot{r}_\gamma + \frac{w_a \dot{r}_\alpha}{r_{\alpha a}}\right) \qquad [16\cdot87b]$$

$$\dot{u}_i = \frac{M}{Z}\left(\frac{2e\Phi_i}{m}\right)^{\frac{1}{2}}\left(w_o' - \frac{w_a}{M}\right) e^{-i\chi_a} \qquad [16\cdot87c]$$

is substituted, an expression for

$$\begin{aligned}\Delta w_i &= e^{i\chi_a}\left(\Delta u_i - i u_i \Delta z_i \left(\frac{eH_i^2}{8mc^2\Phi_i}\right)^{\frac{1}{2}}\right) \\ &= (S_1 + iS_2) w_o'^2 \bar{w}_o' + S_3 w_o' \bar{w}_o' w_a + (S_4 + iS_5) w_o'^2 w_a \\ &\quad + (S_6 - iS_7) \bar{w}_o' w_a^2 + 2(S_6 + iS_7) w_o' w_a \bar{w}_a + S_8 \bar{w}_a w_a^2\end{aligned} \qquad [16\cdot88]$$

is obtained, the coefficients involving integrals over the time parameter. In view of the necessity of subdividing the range of integration to circumvent singularities the expressions for the individual coefficients become too unwieldy to record here. However, the consideration indicates the procedure to be followed in the calculation of the individual aberrations. For example, to determine the spherical aberration, first calculate the relation between t and z by quadrature:

$$t - t_o = \int_{z_o}^{z} \frac{\pm dz}{\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}} \qquad [16\cdot89]$$

making use of the fact that for very small values of $t - t_r$

$$t - t_r = \mp \left(\frac{2m(z_r - z)}{e\Phi_r'}\right)^{\frac{1}{2}} \qquad [16\cdot90]$$

the sign depending on whether t refers to a time before or after reversal

of direction. Then calculate r_α and $\dot{r}_\alpha$ by numerical integration of

$$\ddot{r}_\alpha = -\left(\frac{e\Phi''}{2m} + \frac{e^2H^2}{4m^2c^2}\right) r_\alpha \qquad [16\cdot91]$$

substituting the values of z for given t, as found by the quadrature in Eq. 16·89, in $\Phi''(z)$ and $H(z)$. When the values of the derivatives of Φ and H have been tabulated as function of t, it is possible to obtain Δz by quadrature of Eqs. 16·82, 16·85, and 16·86, with

$$u = \bar{u} = r_\alpha$$

With these values of Δz, $u(t_i)$ is obtained by a second quadrature (Eq. 16·80). All complex terms vanish in the process, as an integration by parts of the terms with $\dot{z}H'''$ and $H''\dot{z}\Delta z$ shows. The aperture defect aberration coefficient is then given by

$$S_8 = \left(\frac{1}{r_{\alpha a}^3}\right)\left\{\Delta u(t_i) + \Delta z(t_i) \cdot \frac{r_{\alpha a}}{Z}\right\} \qquad [16\cdot92]$$

where Z is the distance between the aperture plane and image plane and $r_{\alpha a}$ is the separation from the axis of r_α at the aperture plane, so that $-r_{\alpha a}/Z$ is simply the angle of inclination of the ray in image space.

The case $\Phi_r' = 0$, which was excluded, corresponds to a field in which the point of zero potential on the axis is at the same time a saddle point of the potential. Under these conditions, studied in detail by Recknagel,[8] an infinite number of intermediate images with successively increasing aberrations is formed on approaching the reversal point. At the reversal point itself the motion of the electrons (always in the paraxial approximation) becomes undefined. Electrons of somewhat greater velocity traverse the field, whereas those of smaller velocities and such as are incident on the marginal portions of the mirror lens are reflected. Such a system accordingly does not act as an imaging system for electrons, but as an intensity control for the electron current. As such it is represented by the region about the control grid in any electron gun (Fig. 16·21), when the grid is given

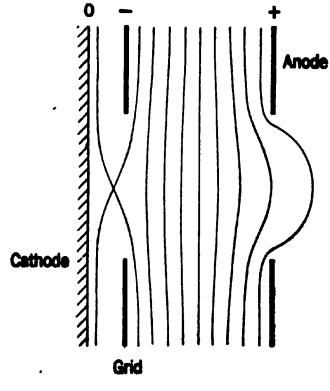


FIG. 16·21. Field Distribution in Front of the Cathode in the Presence of a Negative Grid.

[8] See reference 7.

a negative bias close to the cut-off value. Under normal conditions of operation the presence of space charge in front of the cathode will greatly modify the potential distribution shown.

16·6. Chromatic Aberrations of Electron Lenses. The chromatic aberrations of electron optics are a consequence of nonhomogeneities in the velocities of the image-forming electrons. Since the faster electrons are deflected to a lesser degree by electron lens fields, they are brought to focus at a larger distance from the object. In addition to this, the magnification and rotation of the image will, in general, be different for electrons of different velocities. Formulas for all these effects may be obtained by determining the shift in the intersection of an arbitrary paraxial ray with the Gaussian image plane resulting from a change in the electric potential by a constant amount $\Delta\Phi$.

By introducing once more the function u (Eq. 16·36), satisfying the paraxial ray equation (16·37), and with the particular solutions r_α and r_γ (Eqs. 16·33) vanishing in the object and in the aperture plane, respectively, the general solution can once more be written

$$u = u_o r_\gamma + \frac{u_a r_\alpha}{r_{\alpha a}} \qquad [16\cdot93]$$

u_o and u_a being the values of u in the object and aperture planes.

The solution, Eq. 16·93, is to be compared with the solution of

$$u_1'' + \frac{u_1'\Phi'}{2(\Phi + \Delta\Phi)} + \frac{u_1}{4(\Phi + \Delta\Phi)}\left(\Phi'' + \frac{eH^2}{2mc^2}\right) = 0 \qquad [16\cdot94]$$

which has the same point of intersection and the same slope in the object plane. By writing

$$u_1 = u + \Delta u \qquad [16\cdot95]$$

expanding Eq. 16·94, and retaining only terms of the first order in $\Delta\Phi$ and Δu,

$$\Delta u'' + \Delta u'\frac{\Phi'}{2\Phi} + \Delta u\left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi}\right) = \frac{\Delta\Phi}{\Phi}\left\{\frac{u'\Phi'}{2\Phi} + \frac{u}{4\Phi}\left(\Phi'' + \frac{eH^2}{2mc^2}\right)\right\} = R(z) \qquad [16\cdot96]$$

This may be integrated to yield, for $z = z_i$ (image plane):[9]

$$\Delta u_i = -M\Phi_o^{-\frac{1}{4}}\int_{z_o}^{z_i} R(\zeta)\Phi^{\frac{1}{4}} r_\alpha(\zeta)\, d\zeta \qquad [16\cdot97]$$

$M = \frac{z_i}{z_o}$ being the magnification.

[9] See Adams, reference 4, Formula 8.410.

Now if $w_o' = w_o e^{-i\chi_a}$ represents the coordinates of the object rotated through

$$\chi_a = \int_{z_o}^{z_i} \left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2} dz \qquad [16\cdot98]$$

with respect to the coordinate systems in the aperture and image planes, it is possible to write

$$u = \left(w_o' r_\gamma + \frac{w_a r_\alpha}{r_{\alpha a}}\right) e^{-i\chi_a} \qquad [16\cdot99]$$

so that the shift $\Delta w_i = \Delta x_i + i\Delta y_i$ in the coordinates of the intersection with the image plane is given by

$$\Delta w_i = u_1 e^{i\chi_{1a}} - u_i e^{i\chi_a} = (\Delta u_i + i u_i \Delta\chi_a) e^{i\chi_a} \qquad [16\cdot100]$$

with

$$\Delta\chi_a = -\int_{z_o}^{z_i} \frac{\Delta\Phi}{2\Phi}\left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2} dz \qquad [16\cdot101]$$

If Eqs. 16·97 and 16·101 are substituted in Eq. 16·100, the total chromatic aberration becomes

$$\Delta w_i = \Delta x_i + i\Delta y_i = (C_1 + iC_2) w_o' + C_3 w_a \qquad [16\cdot102]$$

where

$$C_1 = -\frac{M\Delta\Phi}{\Phi_o^{1/2}} \int_{z_o}^{z_i} \left\{\frac{\Phi'}{2\Phi^{3/2}} r_a r_\gamma' + \left[\frac{\Phi''}{4\Phi^{3/2}} + \frac{eH^2}{8mc^2\Phi^{3/2}}\right] r_a r_\gamma\right\} dz \qquad [16\cdot103a]$$

$$C_2 = -\frac{M\Delta\Phi}{2} \int_{z_o}^{z_i} \left(\frac{eH^2}{8mc^2\Phi^3}\right)^{1/2} dz \qquad [16\cdot103b]$$

$$C_3 = -\frac{M\Delta\Phi}{r_{\alpha a}\Phi_o^{1/2}} \int_{z_o}^{z_i} \left\{\frac{\Phi'}{2\Phi^{3/2}} r_a r_\alpha' + \left[\frac{\Phi''}{4\Phi^{3/2}} + \frac{eH^2}{8mc^2\Phi^{3/2}}\right] r_\alpha^2\right\} dz \qquad [16\cdot103c]$$

All chromatic aberration components are thus proportional to the energy differences $e\Delta\Phi$ of the electrons. C_1 corresponds to chromatic difference of magnification, C_2 to chromatic difference in magnetic rotation, C_3 to chromatic difference in image position. If a range of electron velocities corresponding to the energy difference $e\Delta\Phi_{max}$ is present, the first aberration, acting by itself, distorts an image point into a radial line of length $(C_1/\Delta\Phi)\cdot\Delta\Phi_{max}\cdot r_o$, the second into an arc of length $(C_2/\Delta\Phi)\cdot\Delta\Phi_{max}\cdot r_o$, and the third into a circle of confusion with the radius $(C_3/\Delta\Phi)\cdot\Delta\Phi_{max}\cdot r_a$, r_o being the distance of the object point from the axis and r_a the radius of the aperture. It can readily be shown, however, as was first demonstrated by Scherzer,[10] that $C_3/\Delta\Phi$ never

[10] See reference 2.

vanishes or changes sign, so that, in fact, the aberration components C_1 and C_2 cannot be present alone. Since, as an integration by parts shows,

$$\int_{z_o}^{z_i} \frac{\Phi''}{4\Phi^{3/2}} r_\alpha^2 \, dz = -\int_{z_o}^{z_i} \frac{\Phi'}{2\Phi^{3/2}} r_\alpha r_\alpha' \, dz + \int_{z_o}^{z_i} \frac{3\Phi'^2}{8\Phi^{5/2}} r_\alpha^2 \, dz \quad [16\cdot104]$$

C_3 may also be written

$$C_3 = -\frac{M\Delta\Phi}{r_{\alpha a}\Phi_o^{1/2}} \int_{z_o}^{z_i} \left\{ \frac{3\Phi'^2}{8\Phi^{5/2}} + \frac{eH^2}{8mc^2\Phi^{3/2}} \right\} r_\alpha^2 \, dz \quad [16\cdot105]$$

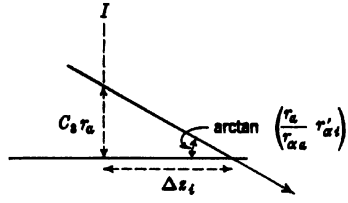


FIG. 16·22. Relation between Longitudinal and Lateral Chromatic Aberration.

Since the terms of the integrand are always positive, the integral is greater than zero for all lens fields. The corresponding shift of the image plane (Fig. 16·22) is given by

$$\Delta z_i = -\frac{C_3 r_a}{\frac{r_a \cdot r'_{\alpha i}}{r_{\alpha a}}} = -\frac{C_3 r_{\alpha a}}{r'_{\alpha i}} \quad [16\cdot106]$$

However, by the theorem of Helmholtz-Lagrange (Eq. 13·27),

$$M = \left(\frac{\Phi_o}{\Phi_i}\right)^{1/2} \frac{r'_{\alpha o}}{r'_{\alpha i}} \quad [16\cdot107]$$

so that

$$\Delta z_i = M^2 \Phi_i^{1/2} \frac{\Delta\Phi}{\Phi_o} \int_{z_o}^{z_i} \left(\frac{3\Phi'^2}{8\Phi^{5/2}} + \frac{eH^2}{8mc^2\Phi^{3/2}} \right) r_\alpha^2 \, dz \quad [16\cdot108]$$

This shift is thus always positive for positive $\Delta\Phi$, that is, larger initial velocities.

No similarly general statement can be made with respect to the remaining chromatic aberration factors C_1 and C_2. Of these C_2 arises only in the presence of a magnetic field and is zero for a rotation-free electron lens formed by an antisymmetric electric and magnetic field.

Glaser[11] has established useful upper limits for the aberration con-

[11] See reference 8.

stant C_3 of purely electric and purely magnetic electron lenses. The upper limit is most simply deduced for purely magnetic lenses. After substituting from the differential equation in Eq. 16·103c, an integration by parts yields

$$C_3 = \frac{M\Delta\Phi}{r_{aa}\Phi}\int_{z_o}^{z_i} r_\alpha'' r_\alpha \, dz = -\frac{M\Delta\Phi}{r_{aa}\Phi}\int_{z_o}^{z_i} r_\alpha'^2 \, dz \qquad [16\cdot109]$$

The differential equation states that r_α'' is always of opposite sign to r_α, so that the path is, throughout, concave toward the axis. If there is no intermediate image r_α must thus rise with continuously decreasing slope

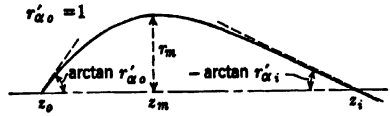


FIG. 16·23. Determination of Upper Limit of Chromatic Aberration for a Magnetic Lens.

from z_o to the point of maximum separation with the abscissa z_m and from there on decline with increasing negative slope toward the image point z_i (Fig. 16·23). Thus

$$|C_3| = \left|\frac{M\Delta\Phi}{r_{aa}\Phi}\right|\left(\int_{z_o}^{z_m} |r'_\alpha| \, dr + \int_{z_i}^{z_m} |r'_\alpha| \, dr\right)$$
$$\leq \left|\frac{M\Delta\Phi}{r_{aa}\Phi}\right| r_m(r'_{ao} - r'_{ai}) \qquad [16\cdot110]$$

and, since $r'_{ai}/r'_{ao} = r'_{ai} = 1/M$,

$$|C_3| \leq \left|\frac{M\Delta\Phi}{r_{aa}\Phi}\right| r_m\left(1 + \frac{1}{|M|}\right) \qquad [16\cdot111]$$

where r_m is the maximum value of r_α, corresponding to the most effective position of the aperture. For the important case of very large magnification $r_m = r_{aa}$ and

$$|C_3| \leq \left|\frac{M\Delta\Phi}{\Phi}\right| \qquad [16\cdot112]$$

For a thin magnetic lens

$$f = \frac{\Phi}{\dfrac{e}{8mc^2}\displaystyle\int_{-\infty}^{\infty} H^2 \, dz}$$

and hence

$$\frac{\Delta f}{f} = \frac{\Delta\Phi}{\Phi} \qquad [16\cdot113]$$

Furthermore, the relations $1/v + 1/u = 1/f$ and $v/u = |M|$ lead to $\Delta v/v = (|M| + 1)\Delta f/f$, so that for a thin lens

$$C_3 = \frac{\Delta r_{ai}}{r_{ao}} = -\frac{\Delta v \cdot r'_{ai}}{r_{ao}} = \frac{|M|\Delta\Phi}{r_{ao}\Phi} r_m \left(1 + \frac{1}{|M|}\right) \qquad [16\cdot114]$$

A comparison of this equality with the general relation in Eq. 16·111 shows that the chromatic aberration of any magnetic lens is less than or equal to the chromatic aberration of a thin magnetic lens of the same focal length. Furthermore, the chromatic aberration of the latter is independent of the field distribution except insofar as it determines the magnitude of the focal length of the lens.

For an electric and magnetic lens the path is not necessarily concave to the axis throughout. This, however, is the case for the function $R = r\Phi^{1/4}$, satisfying

$$R'' = -TR \qquad T = \frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 + \frac{eH^2}{8mc^2\Phi} \qquad [16\cdot115]$$

For a purely electric lens, by Eq. 16·105,

$$C_3 = -\frac{2M\Delta\Phi}{r_{ao}\Phi_o^{1/2}} \int_{z_o}^{z_i} \frac{T}{\Phi} R_a^2\, dz = \frac{2M\Delta\Phi}{r_{ao}\Phi_o^{1/2}} \int_{z_o}^{z_i} \frac{R''_a R_a}{\Phi}\, dz \qquad [16\cdot116]$$

An integration by parts leads to

$$C_3 = -\frac{2M\Delta\Phi}{r_{ao}\Phi_o^{1/2}} \int_{z_o}^{z_i} R'_a d\frac{R_a}{\Phi}$$

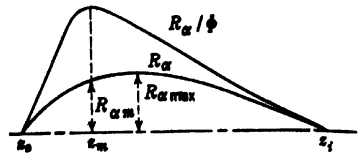


FIG. 16·24. Determination of Upper Limit of Chromatic Aberration for an Electrostatic Lens.

If, now, z_m is the abscissa where R_a/Φ is a maximum (Fig. 16·24),

$$|C_3| \leq \left|\frac{2M\Delta\Phi}{r_{ao}\Phi_o^{1/2}}\right| (|R'_{ao}| + |R'_{ai}|) \frac{R_{am}}{\Phi_m} \qquad [16\cdot117]$$

Since

$$R'_{\alpha o} = r'_{\alpha o}\Phi_o^{1/4} = \Phi_o^{1/4} \qquad R'_{\alpha i} = r'_{\alpha i}\Phi_i^{1/4} = \frac{\Phi_o^{1/4}\Phi_i^{-1/4}}{M}$$

$$\frac{R_{\alpha m}}{\Phi_m} = \frac{r_{\alpha m}}{\Phi_m^{3/4}} \leq \frac{r_{\alpha \max}}{\Phi_{\min}^{3/4}}$$

This inequality results:

$$|C_3| \leq \left|\frac{2M\Delta\Phi}{r_{\alpha a}\Phi_o^{1/4}\Phi_{\min}^{3/4}}\right|\left\{1 + \frac{1}{|M|}\left(\frac{\Phi_o}{\Phi_i}\right)^{1/4}\right\} r_{\alpha \max} \qquad [16\cdot118]$$

For the important case of very great magnification Eq. 16·117 leads to[12]

$$|C_3| \leq \left(\frac{\Phi_i}{\Phi_o}\right)^{1/4}\left|\frac{2M\Delta\Phi}{\Phi_{\min}}\right| \qquad [16\cdot119]$$

whereas for the short electrostatic lens Eq. 16·116 yields

$$C_3 = -2M\Delta\Phi f_i\left(\frac{\Phi_i}{\Phi_o}\right)^{1/2}\int_{z_o}^{z_i}\frac{1}{\Phi}\cdot\frac{3}{16}\cdot\left(\frac{\Phi'}{\Phi}\right)^2 dz \qquad [16\cdot120]$$

where f_i is the image-side focal length of the lens. This depends on the field distribution within the lens and thus does not permit a ready comparison with the general upper limit of the chromatic aberration of an electrostatic lens in Eq. 16·119. For a weak thin lens, however, C_3 is seen to be given by $-2M\Delta\Phi/\Phi$.

16·7. Chromatic Aberrations of the Electron Mirror. To determine the chromatic aberration of an electron mirror it is necessary, again, to turn to the parametric representation of the path equation:

$$\ddot{u} = -\left(\frac{e\Phi''}{2m} + \frac{e^2H^2}{4m^2c^2}\right)u$$

yielding for the change in u due to a change $\Delta\Phi$ in the volt velocity of the electrons

$$\Delta\ddot{u} + \left(\frac{e\Phi''}{2m} + \frac{e^2H^2}{4m^2c^2}\right)\Delta u = -\left(\frac{e\Phi'''}{2m} + \frac{e^2HH'}{2m^2c^2}\right)u\Delta z \qquad [16\cdot121]$$

The latter equation is solved by

$$u(t_i) = \left(\frac{m}{2\Phi_o}\right)^{1/2} M\int_{t_o}^{t_i}\left(\frac{e\Phi'''}{2m} + \frac{e^2HH'}{2m^2c^2}\right)\Delta z\cdot u\cdot r_\alpha\, dt \qquad [16\cdot122]$$

[12] The still closer upper limit $2M\Delta\Phi/\Phi$, where Φ is the object and image space potential of a unipotential lens, does not apply, since the maxima of R and R/Φ need by no means coincide (Fig. 16·24). The formula given by Ramberg, reference 9, footnote 24, is also in error, omitting the factor $r_{\alpha \max}/r_{\alpha a}$.

whereas for $\dot{z} = (2e\Phi/m)^{\frac{1}{2}}$ and, hence,

$$\Delta\dot{z} = \left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\left(\frac{\Delta\Phi}{2\Phi} + \Delta z\,\frac{\Phi'}{2\Phi}\right) \qquad [16\cdot123]$$

the solution becomes, for $t < t_r$ (the time of reversal)

$$\Delta z = \frac{\Delta\Phi}{2}\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\int_{t_o}^{t}\frac{dt}{\Phi} \qquad [16\cdot124]$$

Again, an integration by parts leads to a form useful near $t = t_r$:

$$\begin{aligned}\Delta z &= \left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\left\{-\Delta\Phi\int_{t_r}^{t}\frac{d}{dt}\left(\frac{m}{2e\Phi}\right)^{\frac{1}{2}}\frac{dt}{\Phi'} + \lim_{t\to t_r}\frac{\Delta z_r}{\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}}\right\} \\ &= -\Delta\Phi\left\{\frac{1}{\Phi'} + \left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\int_{t_r}^{t}\frac{\Phi''dt}{\Phi'^2}\right\}\end{aligned} \qquad [16\cdot125]$$

If $\Delta z_2 = \Delta z(t_2)$ where $t_2 > t_r$, Δz is given in the region $t > t_2$ by

$$\Delta z = \frac{\Delta\Phi}{2}\left(\frac{2e\Phi}{m}\right)^{\frac{1}{2}}\int_{t_2}^{t}\frac{dt}{\Phi} - \left(\frac{\Phi}{\Phi_2}\right)^{\frac{1}{2}}\Delta z_2 \qquad [16\cdot126]$$

As for the geometric aberrations, the shift in the image plane is given, finally, by

$$\Delta u_i = \Delta u(t_i) - \Delta z(t_i)\cdot\dot{u}(t_i)\cdot\left(\frac{2e\Phi_i}{m}\right)^{-\frac{1}{2}} \qquad [16\cdot127]$$

and

$$\Delta x_i + i\Delta y_i = \Delta w_i = (\Delta u_i + iu_i\Delta\chi_a)e^{i\chi_a} \qquad [16\cdot128]$$

Since

$$\Delta\chi_a = \int_{t_o}^{t_i}\frac{eH'}{2mc}\Delta z\,dt - \left(\frac{eH_i^2}{8mc^2\Phi_i}\right)^{\frac{1}{2}}\Delta z_i \qquad [16\cdot129]$$

$$\begin{aligned}\Delta w_i &= \left\{\Delta u(t_i) - \Delta z_i\left[\dot{u}_i\left(\frac{2e\Phi_i}{m}\right)^{-\frac{1}{2}} + iu_i\left(\frac{eH_i^2}{8m^2c\Phi_i}\right)^{\frac{1}{2}}\right]\right. \\ &\qquad \left. + iu_i\int_{t_o}^{t_i}\frac{eH'}{2mc}\Delta z\,dt\right\}e^{i\chi_a} \\ &= (C_1 + iC_2)w_o' + C_3 w_a\end{aligned} \qquad [16\cdot130]$$

Here, as on previous occasions, w_o' and w_a represent the coordinates (in complex notation) of the intersections of the ray considered with the object plane and the "aperture plane" respectively; r_α and r_γ are the solutions of the paraxial equation vanishing in the object and the aper-

ture planes, respectively, as defined by Eq. 16·72; and u is given by

$$u = \left(w_o' r_\gamma + \frac{w_a r_\alpha}{r_{\alpha a}}\right) \exp\left(-i \int_{t_o}^{t_i} \frac{eH}{2mc}\, dt\right) \qquad [16\cdot131]$$

As for the geometric aberrations, these rather formidable equations become quite simple for the case of purely electrostatic mirror fields. The complex part (and hence C_2), in particular, vanishes in this case. Furthermore, quite generally, the higher derivatives, Φ''' and H', can be eliminated by integrations by parts from Eq. 16·122 except in the region $t_1 < t < t_2$, that is, except near the point of reversal t_r.

In order to obtain the chromatic aberration on the axis, C_3, u may be set equal to $w_a(r_\alpha/r_{\alpha a})e^{-ix_a}$; the terms with the imaginary factor i in Eq. 16·130 do not contribute to this aberration component (nor to C_1). In many cases the value of C_3 is most conveniently obtained by calculating a number of paraxial rays r_a for electric potential distributions differing by an additive constant $\Delta\Phi$ through the system considered. If the axial position of the image point is plotted against the potential at any fixed point, the slope of the curve yields the value of C_3:

$$-\frac{dz_i}{d\Phi} = \frac{C_3}{\Delta\Phi} \cdot \frac{r_{\alpha a}}{r_{\alpha i}'} \qquad [16\cdot132]$$

16·8. Aberrations of Cathode Lenses. The aberrations of cathode lenses cannot be classified in the same manner as those of ordinary electron lenses and mirrors. This becomes obvious if it is noted that, in the absence of a distribution of initial velocities, the paths of all electrons leaving a given object point (point of the cathode) are identical. Thus, without initial velocities, a sharp image would be produced in every plane normal to the optic axis. It has already been seen in section 13·9, however, that an image plane can be uniquely established if infinitesimal initial lateral velocities are assumed. The image in this plane will normally be distorted. This distortion, however, is the only geometrical aberration present if the initial velocities are taken to be equal to zero.

If the initial velocities are not neglected, the electron paths leaving any object point form pencils (Fig. 16·25) whose width is determined by the maximum value of the initial lateral velocity components. Under the conditions which are prerequisite for the formation of good images with cathode lenses, that is, applied voltages which are very high compared to the initial volt velocities of the electrons, the init al paths of the electrons will be parabolic (corresponding to motion in the nearly uniform field just in front of the cathode), soon forming pencils of very narrow angular aperture. Since all electrons leaving the object point P_o with

zero axial velocity appear, in a plane a distance z from P_o, to diverge from the point P_o', a distance $2z$ from the plane in question along the normal through P_o, irrespective of their initial lateral velocity component

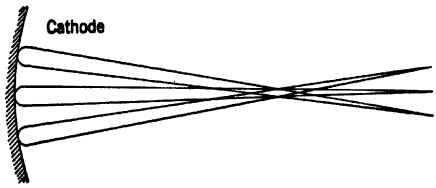


FIG. 16·25. Image-Forming Beams in a Cathode Lens.

(Fig. 16·26), and since, furthermore, the beam aperture is small by the time the electron leaves the nearly uniform field, the position of the image plane is not materially affected by the magnitude of the initial lateral components of velocity. They determine, rather, the aperture angle of the imaging beam. On the other hand, the position of the image plane depends greatly on the initial axial velocity components,

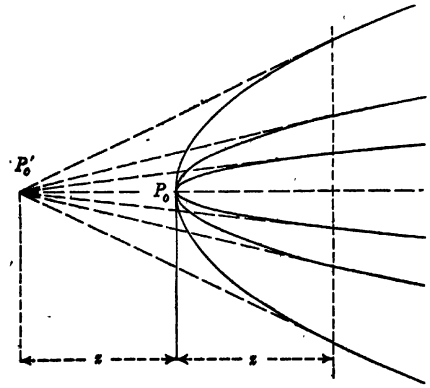


FIG. 16·26. Focusing Effect of Uniform Electric Field on Electrons Leaving Cathode with Zero Axial Velocity.

since these shift the vertex of the parabolas representing the first part of the paths by the amount $\Delta\Phi_z/\Phi_o'$, where $e\Delta\Phi_z$ is the initial kinetic energy of the electron considered associated with motion in the axial direction and $-\Phi_o'$ is the electric field in front of the cathode.

For cathode lenses involving large object fields the image quality may suffer in addition, in the outer portions of the image, from serious defects resulting from the difference in the field conditions surrounding

the principal rays (that is, paths of electrons with zero initial velocity) from the outlying points of the object and those surrounding the optic axis. These defects resemble in all respects the curvature of field and the astigmatism of ordinary lenses and may suitably be described as such. For purely electric fields they are completely determined when the radii of curvature of the *tangential* and *sagittal* image surfaces have been established, provided that the range of lateral initial velocities is known.

In the treatment given below it will be assumed that the cathode is an equipotential surface[13] and may be curved, the radius of curvature on the axis being of necessity $R = 2\Phi_o'/\Phi_o''$, the subscript o designating quantities measured at the cathode. Furthermore, the determination of the aberrations must be preceded by finding the position of the image plane and the magnification of the image near the axis. For the object point on the axis ($C = 0$), the paraxial equation is, once more,

$$r'' = -\frac{\Phi'}{2\Phi}r' - \left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi}\right)r \qquad [16\cdot133]$$

Replacing r with the convergence variable

$$b = -\frac{r'}{r} + \frac{1}{2z} \qquad [16\cdot134]$$

which remains finite at $z = 0$, results in the equation

$$b' = b^2 - b\left(\frac{1}{z} + \frac{\Phi'}{2\Phi}\right) + \frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi} + \frac{1}{2z}\left(\frac{\Phi'}{2\Phi} - \frac{1}{2z}\right) \qquad [16\cdot135]$$

The initial conditions become

$$b_o = \frac{\Phi_o''}{4\Phi_o'} + \frac{eH_o^2}{12mc^2\Phi_o'} \qquad [16\cdot136a]$$

$$b_o' = \frac{2\Phi_o'''}{15\Phi_o'} - \frac{\Phi_o''^2}{40\Phi_o'^2} + \frac{eH_oH_o'}{10mc^2\Phi_o'} - \frac{eH_o^2\Phi_o''}{60mc^2\Phi_o'^2} + \frac{e^2H_o^4}{360m^2c^4\Phi_o'^2} \qquad [16\cdot136b]$$

as may be shown by expanding b, Φ, and H in powers of $z - z_o$ and satisfying Eq. 16·135 for the coefficient of $1/z$ and for the terms independent of z separately. A numerical integration of Eq. 16·135 estab-

[13] It is of course perfectly possible to give the cathode a potential varying in a predetermined manner from the center outward, a procedure which may be utilised to correct aberrations. Normally, however, simpler methods of attaining the same end, for example, giving the cathode a suitable curvature (see Zworykin and Morton, reference 10, and Morton and Ramberg, reference 11), will prove more satisfactory.

lishes the location of the image point z_i, since

$$-\frac{r}{r'} = \frac{2z}{2zb - 1}$$

is the axial separation of the point z,r and the intersection with the axis of the tangent to the curve $r(z)$ at this point (Fig. 16·27). At the same time, a numerical integration of

$$r'' = -\frac{\Phi'}{2\Phi}r' - \left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi} - \frac{eH_o^2}{8mc^2\Phi}\frac{r_o^4}{r^4}\right)r \qquad [16\cdot137]$$

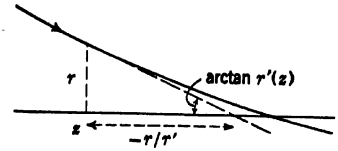


FIG. 16·27. The " Convergence " of an Electron Ray.

the constant C having been determined by placing $\theta_o' = 0$ in Eq. 15·50 with

$$r_o = 1 \qquad r_o' = \frac{\Phi_o''}{2\Phi_o'} \qquad [16\cdot138]$$

yields the magnification of the image and its rotation:

$$\theta_i - \theta_o = \int_{z_o}^{z_i}\left\{\left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2} - \frac{r_o^2}{r^2}\left(\frac{eH_o^2}{8mc^2\Phi}\right)^{1/2}\right\}dz \qquad [16\cdot139]$$

The calculation of the *distortion* of the image is most readily carried out, as for the electron mirror, if a parametric representation of the path is introduced. If it is assumed that an electrostatic field only is present, in the paraxial approximation,

$$\ddot{r} = -\frac{e\Phi''}{2m}r \qquad \dot{z} = \left(\frac{2e\Phi}{m}\right)^{1/2} \qquad [16\cdot140]$$

The first of these equations must be compared with the second-order relation

$$\ddot{r}_s = -\frac{e\Phi_s''}{2m}r_s + \frac{e\Phi^{IV}}{16m}r^3 \qquad [16\cdot141]$$

Subtracting the first-order from the second-order equation leads to a differential equation for $\Delta r = r_s - r$:

$$\Delta\ddot{r} + \frac{e\Phi''}{2m}\Delta r = \frac{e\Phi^{IV}}{16m}r^3 - \frac{e\Phi'''}{2m}\Delta z\cdot r = R(t) \qquad [16\cdot142]$$

Here $r(t)$ in the expression $R(t)$, on the right of Eq. 16·142, is the solution of Eqs. 16·140 with the initial conditions $r(t_0) = 1$, $\dot{r}(t_0) = 0$. The variation with t of Φ'', Φ''', Φ^{IV}, which are given as functions of z, is established by determining $z(t)$ from

$$\begin{aligned} t - t_0 &= \left(\frac{2mz}{e\Phi_0'}\right)^{1/2} && z \leq z_1 \\ t - t_1 &= \int_{z_1}^{z} \frac{dz}{\left(\frac{2e\Phi}{m}\right)^{1/2}} && z > z_1 \end{aligned} \qquad [16·143]$$

where z_1 is sufficiently small that for $z < z_1$ the field in front of the cathode may be regarded as uniform.

Assume now that a second solution of the paraxial equation for r (Eq. 16·140), $\rho(t)$, has been calculated, with the boundary conditions $\rho(t_i) = 0$, $\dot{\rho}(t_i) = 1$. Then, since $r(t_i) = M$, Eq. 16·142 may be integrated to yield:[14]

$$\Delta r(t_i) = -\int_{t_0}^{t_i} R(\tau)\cdot\rho(\tau)\cdot d\tau \qquad [16·144]$$

The distortion itself is given by

$$\Delta r_i = \Delta r(t_i) - \left(\frac{2e\Phi}{m}\right)^{-1/2} \dot{r}_i \Delta z(t_i) \qquad [16·145]$$

For Δz the appropriate equations from the theory of the electron mirror apply. Thus Eqs. 16·85 and 16·86 become, for purely electrostatic cathode lenses, replacing u by r and putting $H \equiv 0$, for t near t_0:

$$\Delta z = \frac{\Phi'' r^2}{4\Phi'} + \frac{m\dot{r}^2}{2e\Phi'} - \left(\frac{2e\Phi}{m}\right)^{1/2} \int_{t_0}^{t} r^2 \left\{\frac{\Phi'''}{4\Phi'} - \frac{\Phi''^2}{4\Phi'^2} - \frac{m\Phi''}{2e\Phi'^2}\frac{\dot{r}^2}{r^2}\right\} dt \qquad [16·146]$$

and for $t > t_1 > t_0$:

$$\Delta z = \left(\frac{\Phi}{\Phi_1}\right)^{1/2} \Delta z_1 - \left(\frac{2e\Phi}{m}\right)^{1/2} \int_{t_1}^{t} \left\{\frac{\Phi'' r^2}{8\Phi} + \frac{m\dot{r}^2}{4e\Phi}\right\} dt \qquad [16·147]$$

As the next step, the *chromatic aberration* of the cathode lens is studied. The shift in the image point resulting from giving the electrons leaving the center of the cathode an initial velocity $\dot{z}_0 = (2e\Delta\Phi_z/m)^{1/2}$ may be

[14] See Adams, reference 4, Formula 8.410.

determined by introducing the convergence variable

$$b_1 = \frac{1}{2\left\{z + \frac{\Delta\Phi_z}{\Phi_o'} - \left[\frac{\Delta\Phi_z}{\Phi_o'}\left(z + \frac{\Delta\Phi_z}{\Phi_o'}\right)\right]^{\frac{1}{2}}\right\}} - \frac{r'}{r} \qquad [16\cdot148]$$

in the paraxial ray equation (with $C = 0$):

$$r'' = -\frac{\Phi'}{2(\Phi + \Delta\Phi_z)}r' - \left\{\frac{\Phi''}{4(\Phi + \Delta\Phi_z)} + \frac{eH^2}{8mc^2(\Phi + \Delta\Phi_z)}\right\}r \qquad [16\cdot149]$$

The first term on the right of Eq. 16·148 represents the convergence $-r'/r$ for an electron traveling through $z = 0$ with the axial velocity

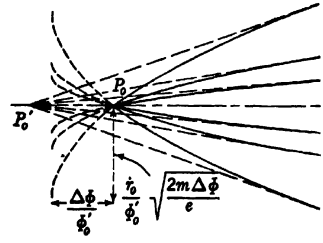


FIG. 16·28. Focusing Action of Uniform Electric Field on an Electron Pencil Diverging from a Point P_o with Uniform Axial Velocity $(2e\Delta\Phi/m)^{\frac{1}{2}}$.

$(2e\Delta\Phi_z/m)^{\frac{1}{2}}$ in the uniform field $-\Phi_o'$, its path being the parabola (see Fig. 16·28):

$$r + \frac{\dot{r}_o}{\Phi_o'}\left(\frac{2m\Delta\Phi_z}{e}\right)^{\frac{1}{2}} = \dot{r}_o\left(\frac{2m}{e\Phi_o'}\left[z + \frac{\Delta\Phi_z}{\Phi_o'}\right]\right)^{\frac{1}{2}} \qquad [16\cdot150]$$

so that

$$\frac{r'}{r} = \frac{\dot{r}}{r}\left(\frac{m}{2e(\Phi_o'z + \Delta\Phi_z)}\right)^{\frac{1}{2}} = \frac{1}{2\left(z + \frac{\Delta\Phi_z}{\Phi_o'}\right)^{\frac{1}{2}}\left[\left(z + \frac{\Delta\Phi_z}{\Phi_o'}\right)^{\frac{1}{2}} - \left(\frac{\Delta\Phi_z}{\Phi_o'}\right)^{\frac{1}{2}}\right]} \qquad [16\cdot151]$$

Hence b_1 is finite at $z = 0$.

Substituting b_1 in Eq. 16·149 and subtracting the equation for b (Eq. 16·135) from the result lead to the following equation for $\Delta b = b_1 - b$:

$$\frac{d\Delta b}{dz} = \Delta b\left(2b - \frac{1}{z} - \frac{\Phi'}{2\Phi}\right) - \left(\frac{\Delta\Phi_z}{\Phi_o'z}\right)^{\frac{1}{2}}\left(\frac{b}{z} + \frac{1}{4z^2} - \frac{\Phi'}{4\Phi_z}\right) = A(z)\Delta b + B(z) \qquad [16\cdot152]$$

Where the terms not containing Δb have been expanded in powers of $\Delta\Phi_z/(\Phi_o' z)$ only the lowest-power members have been retained. The differential equation can be integrated, yielding[15]

$$\Delta b = \Delta b_o \exp\left(\int_o^z A(z')\,dz'\right) + \int_o^z B(z') \exp\left(\int_{z'}^z A(z'')\,dz''\right) dz' \qquad [16\cdot153]$$

The first term may, for all $z \neq 0$, be omitted since the integral in the exponent is negatively infinite, being for small z equal to

$$\lim_{\zeta \to 0}\left(-\frac{3}{2}\log\frac{z}{\zeta}\right)$$

The second term remains finite provided that $B(z)$ does not become infinite more rapidly than in proportion to $1/z^2$ as z approaches 0 — a condition which is fulfilled both by the exact value of $B(z)$ and by the approximate value given by Eq. 16·152.

Since $\Delta b = -\Delta(r'/r)$, the longitudinal chromatic aberration or shift of the image plane is given by

$$\Delta z_i = -\Delta b_a (z_i - z_a)^2 = \frac{\Delta b_a (2z_a)^2}{(2z_a b_a - 1)^2} \qquad [16\cdot154]$$

where the region between z_a and z_i is assumed to be field-free, and the lateral chromatic aberration by

$$\Delta r_i = -\theta_i \Delta z_i \qquad [16\cdot155]$$

θ_i being the aperture angle of the imaging pencil in image space. If the electrons leave the cathode with a maximum lateral velocity component $(e\Delta\Phi_r/m)^{1/2}$, the theorem of Helmholtz-Lagrange yields the value

$$\theta_i = \frac{1}{M}\left(\frac{\Delta\Phi}{\Phi_i}\right)^{1/2} \sin\theta_o = \frac{1}{M}\left(\frac{\Delta\Phi_r}{\Phi_i}\right)^{1/2} \qquad [16\cdot156]$$

Since, by Eqs. 16·153 and 16·152 Δb_o, for small $\Delta\Phi_z$, is proportional to $(\Delta\Phi_z/\Phi_o')^{1/2}$ and since, furthermore, Φ_i is proportional to Φ_o',

$$\Delta r_i = A\frac{(\Delta\Phi_z \Delta\Phi_r)^{1/2}}{\Phi_o'} \qquad [16\cdot157]$$

A being a constant characteristic of the imaging system. It is interesting to note that for very small values of the initial velocities the chromatic aberration of a cathode lens is independent of any magnetic field

[15] See Adams, reference 4, Formula 8.002.

present except insofar as it influences the value of b. The fact, suggested by Eq. 16·157, that the ratio of the initial kinetic energy of the electron to the field at the surface of the cathode is the determining factor for the chromatic aberration is brought out by the examination of some important special cases.

Consider, first, a flat cathode and uniform accelerating field Φ_i/l of length l terminated by an equipotential electric or magnetic short lens (Fig. 16·29). Here the path in the field up to the lens is a parabola such

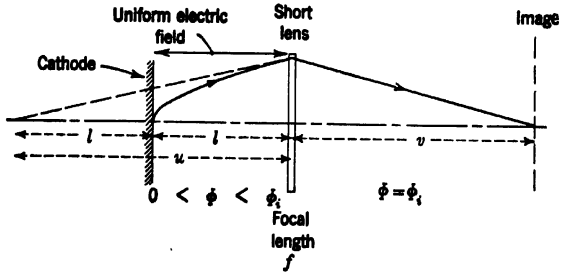


FIG. 16·29. Idealized Cathode Lens.

as is shown in Figs. 16·26 and 16·28. According to Eq. 16·151 the distance of apparent origin of the rays leaving the cathode with axial kinetic energy $e\Delta\Phi_z$ from the lens as they strike the latter is given by

$$\begin{aligned} u = -\frac{r}{r'} &= -2\left\{l + \frac{\Delta\Phi_z}{\Phi_i}l - \left[\frac{\Delta\Phi_z l}{\Phi_i}\left(l + \frac{\Delta\Phi_z l}{\Phi_i}\right)\right]^{1/2}\right\} \\ &= -2l\left\{1 - \left(\frac{\Delta\Phi_z}{\Phi_i}\right)^{1/2} + \cdots\right\} \end{aligned} \qquad [16\cdot158]$$

If v is the image distance and f is the focal length of the short lens,

$$\frac{\Delta v}{v^2} = \frac{\Delta u}{4l^2} + \frac{\Delta f}{f^2} = \frac{1}{2l}\left(\frac{\Delta\Phi_z}{\Phi_i}\right)^{1/2} + \text{const}\,\frac{\Delta\Phi_z}{\Phi_i}\cdots \qquad [16\cdot159]$$

Hence, if only the term of lowest order in $\Delta\Phi_z/\Phi_i$ is retained,

$$\Delta v = M^2 \cdot 2l\left(\frac{\Delta\Phi_z}{\Phi_i}\right)^{1/2} \qquad [16\cdot160]$$

and

$$\Delta r_i = -\theta_i \Delta v = -\frac{1}{M}\left(\frac{\Delta\Phi_r}{\Phi_i}\right)^{1/2} \cdot M^2 \cdot 2l\left(\frac{\Delta\Phi_z}{\Phi_i}\right)^{1/2} = -\frac{2M(\Delta\Phi_r\Delta\Phi_z)^{1/2}}{\Phi_o'} \qquad [16\cdot161]$$

Thus here the constant A in Eq. 16·157 has the value $-2M$.

Exactly the same result is obtained in the quite different case of a

uniform magnetic field superposed on a uniform electrostatic field, the length of the fields being l. Here the magnification is unity. According to Eq. 15·168,

$$\Delta z_i = \frac{2\pi nc}{H}\left(\frac{2m\Delta\Phi_z}{e}\right)^{\frac{1}{2}} \qquad H = \pi nc\left(\frac{2m\Phi_o'}{el}\right)^{\frac{1}{2}} \qquad [16\cdot162]$$

so that

$$\Delta z_i = 2\left(\frac{l\Delta\Phi_z}{\Phi_o'}\right)^{\frac{1}{2}} \qquad [16\cdot163]$$

whereas by Eq. 15·167a

$$\Delta r_i = \left(\frac{2e\Delta\Phi_r}{m}\right)^{\frac{1}{2}}\cdot\Delta t = -\left(\frac{2e\Delta\Phi_r}{m}\right)^{\frac{1}{2}}\cdot\Delta z_i\cdot\left(\frac{2e\Phi_i}{m}\right)^{-\frac{1}{2}}$$

$$= -\frac{2(\Delta\Phi_z\Delta\Phi_r)^{\frac{1}{2}}}{\Phi_o'} \qquad [16\cdot164]$$

Finally an evaluation of Eq. 16·153 for a particular electrostatic cathode lens leads to the value[16]

$$\Delta r_i = -\frac{2M(\Delta\Phi_z\Delta\Phi_r)^{\frac{1}{2}}}{\Phi_o'} \qquad [16\cdot165]$$

within 10 per cent, suggesting a general validity of this formula.

As for other electron lenses, variations in the initial velocities cause with cathode lenses, in addition to a change in position of the image, changes in magnification and rotation. Expressions are readily derived for these changes, starting from the paraxial equation of an electron with the initial kinetic energy $e\Delta\Phi$ associated with motion normal to the cathode:

$$(\Phi + \Delta\Phi)r'' = -\frac{r'\Phi'}{2} - r\left(\frac{\Phi''}{4} + \frac{eH^2}{8mc^2} - \frac{eH_o^2}{8mc_o^2}\frac{r_o^4}{r^4}\right) \qquad [16\cdot166]$$

By subtracting from this the equation for $\Delta\Phi = 0$ and denoting the corresponding change in r by Δr,

$$\Delta r'' + \frac{\Phi'}{2\Phi}\Delta r' + \left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi} + \frac{3eH_o^2}{8mc^2\Phi}\frac{r_o^4}{r^4}\right)\Delta r = R$$

$$R = \frac{\Delta\Phi}{\Phi}\left\{\frac{r'\Phi'}{2\Phi} + r\left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi} - \frac{eH_o^2}{8mc^2\Phi}\frac{r_o^4}{r^4}\right)\right\} \qquad [16\cdot167]$$

$r(z)$ here denotes the solution of Eq. 16·166 with $\Delta\Phi = 0$ and the initial

[16] See Morton and Ramberg, reference 11.

conditions

$$r_o = 1 \qquad r_o' = -\frac{\Phi_o''}{2\Phi_o'}$$
$$r_o'' = -\frac{\Phi_o'''}{6\Phi_o'} + \frac{\Phi_o''^2}{4\Phi_o'^2} + \frac{eH_o^2\Phi_o''}{6mc^2\Phi_o'^2} - \frac{eH_oH_o'}{6mc^2\Phi_o'} \qquad [16\cdot168]$$

Setting $R(z) = 0$, a solution $\rho(z)$ can be obtained (by numerical integration) for Eq. 16·167, which is finite throughout and has the initial values

$$\rho_o = 1 \qquad \rho_o' = -\frac{\Phi_o''}{2\Phi_o'} - \frac{eH_o^2}{mc^2\Phi_o'}$$
$$\rho_o'' = -\frac{\Phi_o'''}{6\Phi_o'} + \frac{\Phi_o''^2}{4\Phi_o'^2} + \frac{eH_o^2\Phi_o''}{6mc^2\Phi_o'^2} + \frac{e^2H_o^4}{3m^2c^4\Phi_o'^2} - \frac{eH_oH_o'}{6mc^2\Phi_o'} \qquad [16\cdot169]$$

With the aid of this solution, the solution of the unreduced Eq. 16·167 may be written[17]

$$\Delta r = \rho \int^z \frac{dz'}{\rho^2(z')\cdot\Phi^{1/2}} \int_0^{z'} \rho(z'')\cdot R(z'')\cdot\Phi^{1/4}\, dz'' \qquad [16\cdot170]$$

It can readily be verified that the above integral converges throughout.

The chromatic difference of rotation is given by

$$\Delta\theta = \int_0^s \left\{\frac{\Delta\Phi}{2\Phi}\left(\frac{eH^2}{8mc^2\Phi}\right)^{1/2} - \left(\frac{\Delta\Phi}{2\Phi} - \frac{2\Delta r}{r}\right)\left(\frac{eH_o^2}{8mc^2\Phi}\right)^{1/2}\frac{r_o^2}{r^2}\right\} dz \qquad [16\cdot171]$$

which is also seen to be convergent. The equations 16·170 and 16·171 show that, as for ordinary electron lenses, the chromatic differences in magnification and rotation are proportional to $\Delta\Phi$.

For the previously considered example of a cathode lens consisting of a uniform field terminated by a unipotential short lens (Fig. 16·29), this change in magnification becomes, as is evident from Fig. 16·30,

$$\Delta r_i = r_o\Delta\left(\frac{v-f}{f}\right) = -r_o\frac{v}{f}\frac{\Delta f}{f} \qquad [16\cdot172]$$

Since $1/f = 1/v + 1/(2l) = (M-1)/(2Ml)$ and $\Delta f/f = K\Delta\Phi/\Phi_i$, where for a purely magnetic short lens $K = 1$,

$$-\Delta r_i = Kr_o(M-1)\frac{\Delta\Phi}{\Phi_i} = K(M-1)\frac{r_o}{l}\frac{\Delta\Phi}{\Phi_o'} \qquad [16\cdot173]$$

The change in rotation, for a purely magnetic short lens, is

$$\Delta\theta = \frac{\Delta\Phi}{2\Phi_i}\left(\frac{e}{8mc^2\Phi_i}\right)^{1/2}\int H\, dz = \frac{\Delta\Phi}{2\Phi_i}(\theta_i - \theta_o) \qquad [16\cdot174]$$

[17] See Adams, reference 4, Formula 8.401.

The change in magnification is seen to be small compared to the effect of the chromatic change in image position as long as the object diameter is smaller than the distance of the object from the lens.

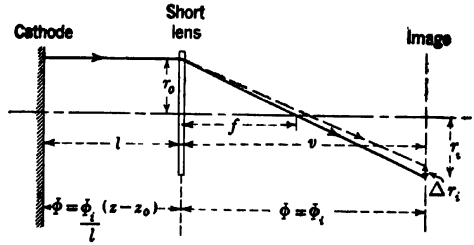


FIG. 16·30. Chromatic Difference of Magnification for Cathode Lens of Fig. 16·29.

Last, the *curvature of field* and the *astigmatism* of cathode lenses will be considered. As in ordinary electron lenses, these defects, if electrostatic fields alone are present, can be characterized by the values of the radii of curvature of the tangential and sagittal image surfaces. On the tangential image surface tangential object elements (for example, circles about the intersection with the axis) are imaged sharply, whereas on the sagittal image surface sagittal object elements (for example, straight lines — *arrows* or *sagittae* — through the origin) appear sharp.

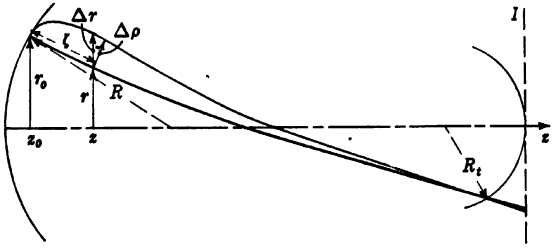


FIG. 16·31. Meridional Ray Pencil Converging at Tangential Image Surface.

In the presence of a magnetic field there are, in general, no such surfaces, the image of a point being throughout, except on the axis, a disk of varying diameter and shape. To simplify the rather complex consideration of these aberrations it will be assumed that no magnetic field is present.

In order to determine the curvature of the tangential image surface, a flat pencil of rays issuing from the object point (Fig. 16·31) and lying in a meridional plane is considered. The exact path equation for any ray in a meridional plane is (Eq. 12·9):

$$r'' = \frac{1 + r'^2}{2\phi}\left\{\frac{\partial \phi}{\partial r} - r'\frac{\partial \phi}{\partial z}\right\} \qquad [16\cdot175]$$

Denote the solution of Eq. 16·175 for the principal ray (electron with zero initial velocity) simply by r, and consider any neighboring ray $r + \Delta r$ in the meridional plane. By Eq. 16·175 the variation of the small quantity Δr with z is given by:

$$\Delta r'' = \frac{\Delta r}{2\phi}(1 + r'^2)\left\{\frac{\partial^2\phi}{\partial r^2} - \frac{1}{\phi}\left(\frac{\partial\phi}{\partial r}\right)^2 + \frac{r'}{\phi}\frac{\partial\phi}{\partial r}\frac{\partial\phi}{\partial z} - r'\frac{\partial^2\phi}{\partial r\partial z}\right\} + \frac{\Delta r'}{2\phi}\left\{2r'\frac{\partial\phi}{\partial r} - (1 + 3r'^2)\frac{\partial\phi}{\partial z}\right\} \qquad [16\cdot176]$$

To bring the equation for the deviation from the principal ray into a form strictly comparable to that for the deviation of a paraxial ray leaving the center of the object from the optic axis, it is advisable to introduce two new variables: the distance of the ray from the principal ray in a direction normal to the latter:

$$\Delta\rho = \frac{\Delta r}{(1 + r'^2)^{1/2}} \qquad [16\cdot177]$$

and the distance ζ along the principal ray:

$$\zeta = \int_{z_o}^{z} (1 + r'^2)^{1/2}\, dz \qquad [16\cdot178]$$

With these variables Eq. 16·176 becomes, with terms of higher than the second order in r and r' omitted:

$$\begin{aligned}
\frac{d^2\Delta\rho}{d\zeta^2} &= -\left(\frac{\Phi'}{2\Phi} - \alpha_t\right)\frac{d\Delta\rho}{d\zeta} - \left(\frac{\Phi''}{4\Phi} - \beta_t\right)\Delta\rho \\
\alpha_t &= r^2\left(\frac{\Phi'''}{8\Phi} - \frac{\Phi'\Phi''}{8\Phi^2}\right) + \frac{rr'\Phi''}{4\Phi} + \frac{r'^2\Phi'}{4\Phi} \\
&\quad - \left(\frac{r_o^2}{2R} - \int_0^z \frac{r'^2}{2}\, dz\right)\left(\frac{\Phi''}{2\Phi} - \frac{\Phi'^2}{2\Phi^2}\right) \\
\beta_t &= r^2\left(\frac{3\Phi^{IV}}{32\Phi} - \frac{\Phi''^2}{4\Phi^2}\right) + rr'\left(\frac{\Phi'''}{2\Phi} - \frac{3\Phi'\Phi''}{4\Phi^2}\right) \\
&\quad + r'^2\left(\frac{3\Phi''}{4\Phi} - \frac{3\Phi'^2}{4\Phi^2}\right) \\
&\quad - \left(\frac{r_o^2}{2R} - \int_0^z \frac{r'^2}{2}\, dz\right)\left(\frac{\Phi'''}{4\Phi} - \frac{\Phi'\Phi''}{4\Phi^2}\right).
\end{aligned} \qquad [16\cdot179]$$

In this expression Φ signifies $\Phi(\zeta)$, not $\Phi(z(\zeta))$. $R = 2\Phi_o'/\Phi_o''$ is the radius of curvature of the cathode.

As for the paraxial equation, the convergence variable

$$b_t = \frac{1}{2\zeta} - \frac{\Delta\rho'}{\Delta\rho} \qquad [16\cdot180]$$

may be introduced, leading to

$$b_t' = b_t^2 + b_t\left(-\frac{\Phi'}{2\Phi} - \frac{1}{\zeta} + \alpha_t\right) + \frac{\Phi'}{4\Phi\zeta} - \frac{1}{4\zeta^2} - \frac{\alpha_t}{2\zeta} + \frac{\Phi''}{4\Phi} - \beta_t \qquad [16\cdot181]$$

By subtracting from this the paraxial equation

$$b' = b^2 + b\left(-\frac{\Phi'}{2\Phi} - \frac{1}{\zeta}\right) + \frac{\Phi'}{4\Phi\zeta} - \frac{1}{4\zeta^2} + \frac{\Phi''}{4\Phi} \qquad [16\cdot182]$$

the differential equation for $\Delta b_t = b_t - b$ is obtained:

$$\begin{aligned} \Delta b_t' &= A_t \Delta b_t + B_t \qquad A_t = 2b - \frac{\Phi'}{2\Phi} - \frac{1}{\zeta} \\ B_t &= \alpha_t\left(b - \frac{1}{2\zeta}\right) - \beta_t \end{aligned} \qquad [16\cdot183]$$

As in the treatment of the chromatic aberration of cathode lenses, this equation is solved by

$$\Delta b_t = \int_0^z B_t(z')\exp\left(\int_{z'}^z A_t(z'')\,dz''\right)dz' \qquad [16\cdot184]$$

Since for very small z the exponential factor is proportional to $(z'/\zeta)^{3/2}$ and as $B_t(z')$ does not go to infinity at $z' = 0$ more rapidly than $1/z'^2$, the integral for Δb_t is convergent.

The quantity Δb_t is simply related to the radius of curvature of the tangential image surface. Assume, for convenience, that for $z_1 < z < z_i$ the electric field is zero. Then the abscissa of the point of convergence of the ray is given by

$$z_i' = \int^{z_1} \frac{d\zeta}{(1 + r'^2)^{1/2}} + \frac{r_o^2}{2R} + \frac{1}{(1 + r_i'^2)^{1/2}} \frac{2z_1}{2z_1 b_{t1} - 1} \qquad [16\cdot185]$$

and the distance between the plane of convergence and the image plane by

$$\begin{aligned} \Delta z_i = z_i' - z_1 - \frac{2z_1}{2z_1 b_1 - 1} &= -\int_0^{z_1} \frac{r'^2}{2}\,dz + \frac{r_o^2}{2R} \\ - \frac{4z_1^2 \Delta b_{t1}}{(2z_1 b_1 - 1)^2} &= C_t r_o^2 \end{aligned} \qquad [16\cdot186]$$

C_t is here a constant depending only on the field distribution between the cathode object and the image. If now the radius of curvature of the tangential image surface is R_t,

$$\frac{1}{R_t} = \frac{2\Delta z_i}{r_i^2} = \frac{2C_t}{M^2} \qquad [16\cdot187]$$

Next it is necessary to determine the radius of curvature of the sagittal image surface. For this purpose it is convenient to introduce rectangu-

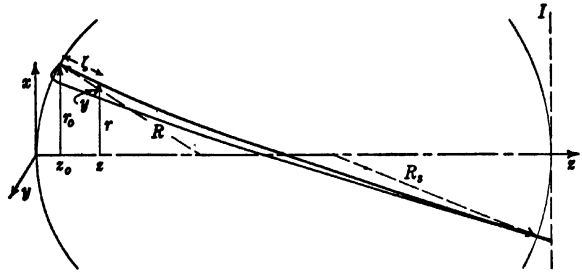


FIG. 16·32. Tangential Ray Pencil Converging at Sagittal Image Surface.

lar coordinates, x and z lying in the meridional plane and y normal to the same (Fig. 16·32). The Euler equations of Fermat's law take the form

$$\frac{\partial F}{\partial x} - \frac{d}{dz}\frac{\partial F}{\partial x'} = 0 \qquad \frac{\partial F}{\partial y} - \frac{d}{dz}\frac{\partial F}{\partial y'} = 0$$
$$F = [\phi(1 + x'^2 + y'^2)]^{1/2} \qquad [16\cdot188]$$

Explicitly, the first equation becomes

$$x'' = \frac{x'y'y''}{1+y'^2} - \frac{x'}{2\phi}\frac{1+x'^2+y'^2}{1+y'^2}\left(\frac{\partial\phi}{\partial r}\frac{xx'+yy'}{r} + \frac{\partial\phi}{\partial z}\right) + \frac{x}{2r\phi}\frac{(1+x'^2+y'^2)^2}{1+y'^2}\frac{\partial\phi}{\partial r} \qquad [16\cdot189]$$

The second equation is obtained from this by interchanging x and y throughout. If y is assumed to be infinitesimal, $x \to r$ and

$$r'' = \frac{1+r'^2}{2\phi}\left(\frac{\partial\phi}{\partial r} - r'\frac{\partial\phi}{\partial z}\right) \qquad [16\cdot190a]$$

$$y'' = \frac{1+r'^2}{2\phi}\left(\frac{y}{r}\frac{\partial\phi}{\partial r} - y'\frac{\partial\phi}{\partial z}\right) \qquad [16\cdot190b]$$

or, by omitting terms of higher order than the second in r and r',

$$y'' = y'\left(-\frac{\Phi'}{2\Phi} - \frac{r'^2\Phi'}{2\Phi} + r^2\left[\frac{\Phi'''}{8\Phi} - \frac{\Phi'\Phi''}{8\Phi^2}\right]\right) + y\left(-\frac{\Phi''}{4\Phi} - \frac{r'^2\Phi''}{4\Phi} + r^2\left[\frac{\Phi^{IV}}{32\Phi} - \frac{\Phi''^2}{16\Phi^2}\right]\right) \quad (16\cdot191]$$

Introducing, as before, ζ (Eq. 16·178) as independent variable and putting

$$b_s = \frac{1}{2\zeta} - \frac{y'}{y} \quad [16\cdot192]$$

lead, in a manner similar to that for the meridional pencil, to

$$\Delta b_s' = A_s \Delta b_s + B_s \quad \text{with} \quad \Delta b_s = b_s - b \quad [16\cdot193]$$

and

$$A_s = 2b - \frac{\Phi'}{2\Phi} - \frac{1}{\zeta} \qquad B_s = \left(b - \frac{1}{2\zeta}\right)\alpha_t - r^2\left(\frac{\Phi^{IV}}{32\Phi} - \frac{\Phi''^2}{16\Phi^2}\right) + \left(\frac{r_o^2}{2R} - \int_0^s \frac{r'^2}{2}\,dz\right)\left(\frac{\Phi'''}{4\Phi} - \frac{\Phi'\Phi''}{4\Phi^2}\right)$$

Here α_t is given by Eq. 16·179. Equation 16·193 is solved by

$$\Delta b_{s_1} = \int_0^{s_1} B_s(z') \exp\left(\int_{z'}^{s_1} A_s(z'')\,dz''\right) dz' \quad [16\cdot194]$$

If, again, the electric field is zero beyond $z = z_1$, the radius of curvature of the sagittal image surface, R_s, is given by

$$\frac{1}{R_s} = \frac{2C_s}{M^2} = \frac{2}{M^2}\left(-\int_0^{s_1} \frac{r'^2}{2r_o^2}\,dz + \frac{1}{2R} - \frac{4z_1^2\Delta b_{s1}}{r_o^2(2z_1b_1 - 1)^2}\right) \quad [16\cdot195]$$

16·9. Space-Charge Effects. Space charge has been used repeatedly for the concentration of electron beams[18] and even for the formation of electron images.[19] In both these cases space charge consists of positive-ion concentration in and near an electron beam passing through an incompletely evacuated space. The electrons, in moving through such a region, ionize gas atoms along their path. While the electrons ejected from the atoms leave the region of ionization with great velocity, the positive ions, owing to their much greater mass, remain behind and form a positive space-charge column which exceeds the negative space charge of the beam itself. Thus under favorable circumstances electrons

[18] See Johnson, reference 12.

[19] See v. Borries and Ruska, reference 13.

diverging from the axis of the beam will be drawn back toward it. This is the origin of the luminous electron " thread " or nodal beams, which form a convenient means for demonstrating the action of deflecting fields on charged bodies. The most important use made of such gas-concentrated beams is probably Johnson's low-voltage cathode-ray tube, which relies entirely on this type of beam concentration. The formation of images by the field of the gas discharge between electrodes of a three-element electrostatic lens configuration has been demonstrated by von Borries and Ruska.

Here, however, space charge is to be treated purely as a disturbing phenomenon. Furthermore, it is assumed that the vacuum in the region of travel of the electrons is sufficiently perfect that no appreciable number of positive ions is formed. Thus only the negative space charge of the electron beam itself need be considered.

The effect of this space charge on the paths of the individual electrons is twofold: On the one hand, it alters the potential distribution along the axis; on the other, it changes the dependence of the potential distribution off the axis on that along the axis. For, assume the presence of an axially symmetric space-charge distribution:

$$\rho = \rho_o + \rho_1 r^2 + \rho_2 r^4 \cdots \qquad [16\cdot196]$$

where $\rho_o, \rho_1, \rho_2 \cdots$ are functions of z. Then the potential

$$\phi = \Phi + ar^2 + br^4 \cdots \qquad [16\cdot197]$$

where, again, Φ, a, $b, \cdots$ are functions of z, must satisfy Poisson's equation:

$$\Delta^2\phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) + \frac{\partial^2\phi}{\partial z^2} = -4\pi\rho \qquad [16\cdot198]$$

Substituting the expansions 16·196 and 16·197 leads to the following equations for the coefficients of the several powers of r:

$$\begin{aligned} 4a + \Phi'' &= -4\pi\rho_o \\ 16b + a'' &= -4\pi\rho_1, \text{ etc.} \end{aligned} \qquad [16\cdot199]$$

whence

$$\begin{aligned} a &= -\frac{\Phi''}{4} - \pi\rho_o \\ b &= \frac{\Phi^{IV}}{64} + \frac{\pi\rho_o''}{16} - \frac{\pi\rho_1}{4}, \text{ etc.} \end{aligned} \qquad [16\cdot200]$$

Accordingly the paraxial ray equation in the presence of axially sym-

metric space charge takes on the form

$$u'' + \frac{u'\Phi'}{2\Phi} + \left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi} + \frac{\pi\rho_o}{\Phi}\right)u = 0 \qquad [16\cdot201]$$

where, as on previous occasions, $u = r\exp\left\{i\left[\theta - \int_{z_o}^{z}\left(\frac{eH^2}{8mc^2\Phi}\right)^{\frac{1}{2}}dz\right]\right\}$

If $\Delta\Phi$ is the change in potential on the axis due to the presence of space charge, the space-charge aberration may thus, in principle, be determined, in a manner similar to the other aberrations, by the solution of

$$\Delta u'' + \frac{\Delta u'\Phi'}{2\Phi} + \Delta u\left(\frac{\Phi''}{4\Phi} + \frac{eH^2}{8mc^2\Phi}\right) \qquad [16\cdot202]$$

$$= -u'\left(\frac{\Delta\Phi'}{2\Phi} - \frac{\Phi'\Delta\Phi}{2\Phi^2}\right) - u\left(\frac{\Delta\Phi''}{4\Phi} - \frac{\Phi''\Delta\Phi}{4\Phi^2} - \frac{eH^2\Delta\Phi}{8mc^2\Phi^2} + \frac{\pi\rho_o}{\Phi}\right)$$

Here $\Delta\Phi(z)$ is to be found from the space-charge distribution in the beam and from its electrical images in the surrounding conductors by the method outlined in section 11·12.

Very frequently, however, all that is required is a check to see whether, under actual conditions, space-charge effects are negligible. Thus, for a beam of uniform cross section along the axis in a region which is field-free except for the presence of the beam $\Phi'' = \Phi' = 0$ and, hence, confining attention to rays in a meridional plane,

$$r'' + \pi\rho_o \cdot \frac{r}{\Phi} = 0 \qquad [16\cdot203]$$

For $r_o = 0$ at $z = 0$ this is solved by

$$r = r_o'\left(-\frac{\pi\rho_o}{\Phi}\right)^{-\frac{1}{2}}\sinh\left[\left(-\frac{\pi\rho_o}{\Phi}\right)^{\frac{1}{2}}z\right] \qquad [16\cdot204]$$

On the other hand, for a ray which is parallel to the axis and a distance r_o from it at $z = 0$,

$$r = r_o\cosh\left[\left(-\frac{\pi\rho_o}{\Phi}\right)^{\frac{1}{2}}z\right] \qquad [16\cdot205]$$

Since

$$\rho_o = -\frac{i}{A}\left(\frac{2e\Phi}{m}\right)^{-\frac{1}{2}} \qquad [16\cdot206]$$

Eq. 16·205 may also be written

$$\begin{aligned} r &= r_o \cosh\left[2\left(\frac{2e}{m}\right)^{-\frac{1}{4}} i^{\frac{1}{4}}\Phi^{-\frac{3}{4}}\frac{z}{D}\right] \\ &= r_o \cosh\left[246.4 i^{\frac{1}{4}}\Phi^{-\frac{3}{4}}\frac{z}{D}\right] \end{aligned} \qquad [16\cdot207]$$

Here A is the cross-section area of the beam, i the beam current in amperes, Φ the accelerating voltage in volts, and D the diameter of the beam. For the small deflections considered the hyperbolic cosine may be expanded, leading to

$$\Delta r = r - r_o = 3.04\cdot10^4 i\Phi^{-\frac{3}{2}}\left(\frac{z}{D}\right)^2 r_o \qquad [16\cdot208]$$

Thus the space-charge deflection of the initially parallel ray is proportional to the current density, the square of the path length, and inversely proportional to the $\frac{3}{2}$ power of the voltage. The shift in the apparent point of origin in the plane $z = 0$ is given by

$$\delta = -r'z + \Delta r = -\Delta r \qquad [16\cdot209]$$

that is, by the same quantity as the deflection of the ray. As an example, assume that $z = 1$ centimeter, $\Phi = 40{,}000$ volts, $D = 10^{-2}$ centimeter, $r_o = 10^{-3}$ centimeter, $i = 10^{-6}$ ampere. Then

$$\Delta r = 4\cdot10^{-8} \text{ cm}$$

For a ray leaving the center of the beam at an inclination r_o' in the plane $z = 0$, the deviation of the apparent point of origin is correspondingly small:

$$\delta = \frac{1}{3} z r_o' \cdot \frac{\pi\rho_o z^2}{\Phi} = -2.03\cdot10^4 z r_o' i\Phi^{-\frac{3}{2}}\left(\frac{z}{D}\right)^2 \qquad [16\cdot210]$$

which for $r_o' = 10^{-2}$ and otherwise the same data as in the previous example leads to

$$\delta = -2.7\cdot10^{-7} \text{ cm}$$

For a more exhaustive treatment of the effect of space charge on the beam diameter, for diverging and converging as well as parallel beams, the reader is referred to the work of Watson, von Borries and Dosse, and Haeff.[20]

In the above consideration two effects, that of charges induced by the beam on surrounding conductors and that of the magnetic field associ-

[20] See Watson, reference 14, v. Borries and Dosse, reference 15, and Haeff, reference 16.

ated with the beam current, have been neglected. Since both of these, however, reduce the space-charge effect, the estimates obtained by Eqs. 16·208 and 16·210 will invariably be too large, or, in other words, on the safe side. The relative magnitude of the repulsive action of the electrostatic field of the beam and the attracting action of its magnetic field can readily be estimated. Thus the electrostatic field force acting on an electron a distance r from a line of charge with λ electrostatic units per centimeter is $2e\lambda/a$. If the charge λ is the charge of a moving beam, this may be written $2ie/(av)$, where i is the beam current in e.s.u. and v the velocity of the charges constituting the beam. The magnetic field, on the other hand, is $2i/(ac)$, where c is the velocity of light, and the consequent force on the electron is $-ehv/c = -2iev/(ac^2)$, electrostatic units being used throughout. Thus the attractive force of the magnetic field is to the repulsive force of the electrostatic field as the square of the ratio of the electron velocity to the velocity of light.

For long, high-current beams, such as are employed in television projection tubes, the effect of space charge is considerable and may result in a change in the conditions of optical focus with the intensity of the beam. For example, on the basis of Eq. 16·208, if the beam is 10 centimeters long and carries 1 milliampere distributed over a cross section of 0.2 millimeter diameter, a ray proceeding initially parallel to the beam direction along its outer edge would, for a volt velocity of 40,000 volts, be deflected by an amount $\Delta r = 0.1$ millimeter, that is, an amount equal to the beam radius. In practice, conditions are rarely as severe as this. Except in the immediate proximity of the object and image the beam will have a much greater diameter than that indicated (0.2 millimeter) and space-charge forces will be correspondingly less.

16·10. Instabilities, Disturbances, and Misalignments. Even if the elements of a particular electron-optical system are in principle capable of producing a perfect image, they will not do so unless both their construction and alignment are perfect and unless the voltages — in the case of electrostatic lenses — or both the voltages and coil currents — in the case of magnetic lenses — are perfectly stable during focusing and observation or exposure. Finally, disturbing fields must be entirely absent.

It is relatively easy to determine the effects of voltage and of current fluctuations as well as of disturbing external fields. The effect of *fluctuations in the accelerating voltage* (with lens voltages with respect to the anode as well as magnetic lens fields assumed to remain constant) is given directly by the formulas for the chromatic aberration. However, if, as is commonly the case for purely electrostatic lenses, the lens electrode voltages are all derived from the same voltage source, fluctua-

tions in the latter will not affect the image formed by the system, at least as long as the electron velocities involved are not comparable with the velocity of light and if the reactive impedance components of the electrodes and the leads joining them to the source remain negligible. If, however, as shown in Fig. 16·33, a resistance is used as voltage divider, thermal variations in the resistance may give rise to fluctuations in the

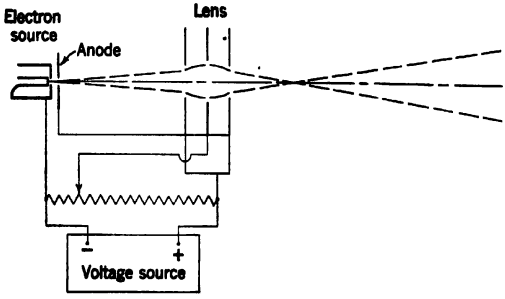


FIG. 16·33. Three-Element Electrostatic Lens Operated with Potential Divider.

lens properties. For this reason the operation of electrostatic lenses involving only electrodes at cathode and anode voltage, making a voltage divider unnecessary, is particularly favorable.

In the case of purely magnetic lenses, the effect of *fluctuations in the magnetic field* may also be simply related to the chromatic aberration. The varying of both the magnetic field and the accelerating potential in the paraxial equation

$$u'' = -\frac{eH^2}{8mc^2\Phi}u \qquad [16\cdot211]$$

leads to

$$u'' = -\frac{eH^2}{8mc^2\Phi}u\left(1 + \frac{2\Delta H}{H} - \frac{\Delta\Phi}{\Phi}\right) \qquad [16\cdot212]$$

If it is assumed that the fluctuation in the magnetic field is due to variations in the coil current i and, furthermore, that either the magnetic lens does not utilize iron or that its iron is quite unsaturated,

$$\frac{\Delta H}{H} = \frac{\Delta i}{i} \qquad [16\cdot213]$$

is a constant along the length of the lens. Thus for

$$\frac{2\Delta i}{i} = \frac{\Delta\Phi}{\Phi} \qquad [16\cdot214]$$

the electron paths are, by Eq. 16·212, unaffected by the simultaneous fluctuations in lens current and accelerating voltage. A lens current fluctuation Δi has the same effect on the ray paths as the accelerating voltage fluctuation $-2\Phi\Delta i/i$. Hence, by Eqs. 16·102 and 16·103, the aberration resulting from a fluctuation of the lens current Δi of a purely magnetic lens may be written

$$\Delta x_i + i\Delta y_i = -2\left(\frac{\Delta i}{i}\right)\left(\frac{\Delta\Phi}{\Phi}\right)^{-1}\{(C_1+iC_2)(x_o'+iy_o') + C_3(x_a+iy_a)\} \qquad [16\cdot215]$$

i denoting the lens current and the imaginary quantity $(-1)^{\frac{1}{2}}$, respectively. Equations 16·103, with $\Phi' \equiv 0$, give the values of C_1, C_2, and C_3.

Disturbing fields may be either electric or magnetic. However, since any thin metallic shield screens off electric fields perfectly, magnetic fields alone need be considered here.

It is always possible to resolve a magnetic field into a component parallel to the optic axis and one normal to it. Since the direction of motion of the electrons deviates but little from that of the axis and since the force action of a magnetic field on electrons is proportional to their

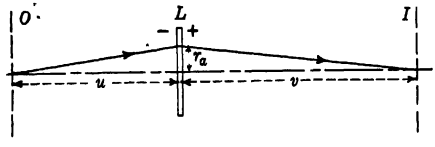


FIG. 16·34. Effect of Disturbing Fields on Imaging by Thin Lens; Significance of Symbols.

component of velocity normal to the field, it is evident that the component normal to the axis will prove more disturbing. Both effects will be considered for a thin lens with object distance u and image distance v (Fig. 16·34).

The effect of the longitudinal field component can be determined readily by dividing the system into three regions: object space ($z < u$), the region occupied by the lens ($z \cong u$), and image space ($u < z < u + v$). Paraxial rays can be traced through the first and third region with the aid of the equations for electron motion in uniformly magnetic fields given in section 15·2.

Thus, for a ray starting from the axial object point with an inclina-

tion r'_o,

$$r_a = \frac{r'_o}{A}\sin Au \qquad r'_a = r'_o \cos Au \tag{16·216}$$

with

$$A = \left(\frac{eh_z^2}{8mc^2\Phi}\right)^{1/2} = \frac{0.148h_z}{\Phi^{1/2}} \tag{16·217}$$

h_z being the disturbing field component along the z-axis.

At the lens,

$$-\frac{r'_{a+}}{r_a} = -\frac{r'_{a-}}{r_a} + \frac{1}{u} + \frac{1}{v} = -\frac{A}{\tan Au} + \frac{1}{u} + \frac{1}{v} \cong \frac{1}{v} + \frac{A^2u}{3} \tag{16·218}$$

u and v being taken as positive. Thus, in the image plane,

$$\begin{aligned}\Delta r_i = r_i &= r_a\left\{\cos Av - \left(\frac{1}{v} + \frac{A^2u}{3}\right)\frac{\sin Av}{A}\right\} \\ &= -\frac{r'_oA^2}{3}uv(v+u) = -\frac{r_a}{3}M^2\left(1+\frac{1}{M}\right)A^2u^2\end{aligned} \tag{16·219}$$

Here r_a denotes the aperture radius of the lens and M the magnification.

For a paraxial ray incident parallel to the axis from an object point a distance r_o from the axis, however,

$$r_a = r_o \cos Au \qquad r'_a = -r_oA \sin Au \tag{16·220}$$

$$-\frac{r'_{a+}}{r_a} = A\tan Au + \frac{1}{u} + \frac{1}{v} = \frac{1}{u} + \frac{1}{v} + A^2u \tag{16·221}$$

and

$$\begin{aligned}\Delta r_i &= r_o\cos Au\left\{\cos Av - \left(\frac{1}{u}+\frac{1}{v}+A^2u\right)\frac{\sin Av}{A}\right\} + r_o\frac{v}{u} \\ &= \frac{r_o}{6}A^2M^2u^2\left(1-\frac{3}{M}\right)\left(1+\frac{1}{M}\right)\end{aligned} \tag{16·222}$$

In addition to this there is a lateral shift

$$r_i\Delta\theta_i = r_oMA(u+v) = r_oM^2Au\left(1+\frac{1}{M}\right) \tag{16·223}$$

For example, for a point on the axis,

$$\Delta r_i = -\frac{0.022}{3}M^2u^2\left(1+\frac{1}{M}\right)\frac{h_z^2}{\Phi}r_a \text{ cm} \tag{16·224}$$

becomes, with $u = 1$ centimeter, $M = 100$, $r_a = 3\cdot10^{-3}$ centimeter, $h_z = 1$ gauss, $\Phi = 50{,}000$ volt, $\Delta r_i = -4.4\cdot10^{-6}$ centimeter, whereas,

under the same conditions, for a marginal point with $r_o = 10^{-3}$ centimeter, there is, in addition, a distortion $\Delta r_i = 7.3 \cdot 10^{-5}$ centimeter, $r_i \Delta\theta_i = 6.6 \cdot 10^{-3}$ centimeter. Thus the rotational distortion is easily the most serious consequence of a fluctuating magnetic field along the axis of the instrument. A superposed stationary axial field is of course inconsequential, since it may be regarded as part of the electron lens system.

The effect of a transversal disturbing field is, as mentioned above, greater than that of a longitudinal field. The radius of curvature R of the path of an electron traveling with the velocity $(2e\Phi/m)^{1/2}$ in the transverse field h_x is (according to Eq. 15·20) given by

$$\frac{1}{R} = \left(\frac{eh_x^2}{2mc^2\Phi}\right)^{1/2} = \frac{0.296h_x}{\Phi^{1/2}} \qquad [16\cdot225]$$

In object and image space the electron paths are circles with this radius. For an electron starting from $z_o = 0$, $r_o = 0$ with an inclination r_o', the path in object space becomes

$$\left[r - \frac{R}{(1 + r_o'^2)^{1/2}}\right]^2 + \left[z + \frac{r_o' R}{(1 + r_o'^2)^{1/2}}\right]^2 = R^2 \qquad [16\cdot226]$$

If quantities of higher than the first order in r_o' and higher than the second order in u/R are neglected, it follows that just before the lens

$$r_a = \frac{u^2}{2R} + r_o' u \qquad r_{a-}' = \frac{u + r_o' R}{R - \dfrac{u^2}{2R} - r_o' u} \qquad [16\cdot227]$$

and hence

$$r_{a+}' = r_{a-}' - r_a\left(\frac{1}{u} + \frac{1}{v}\right) = \frac{u}{R} + r_o' - \left(\frac{u^2}{2R} + r_o' u\right)\left(\frac{1}{u} + \frac{1}{v}\right) \qquad [16\cdot228]$$

Similarly, in image space the path is given by

$$\left[r - r_a - \frac{R}{(1 + r_{a+}'^2)^{1/2}}\right]^2 + \left[z - u + \frac{r_{a+}' R}{(1 + r_{a+}'^2)^{1/2}}\right]^2 = R^2 \qquad [16\cdot229]$$

and, for $z_i = u + v$,

$$\Delta r_i = \frac{v^2 + uv}{2R} = \frac{M^2 u^2}{2R}\left(1 + \frac{1}{M}\right) \qquad [16\cdot230]$$

For $u = 1$ centimeter, $M = 100$, $h_x = 1$ gauss, $\Phi = 50{,}000$ volts, as before, $\Delta r_i = 6.6$ centimeters. The deviation for the transverse field is thus of an entirely different order of magnitude from that for the axial field. The fact that r_o' does not appear in Eq. 16·230 justifies the con-

clusion that, in this order of approximation, the deflection is not dependent on the orientation of the ray at the object point. A calculation for off-axis object points reveals that the shift of the image point is given by Eq. 16·230 in this case also. A comparison of the figures for the deflecting effect of the normal and axial magnetic field components suggests that any magnetic shielding[21] which adequately reduces the transversal component will do at least as much for the axial component.

An obvious effect of any *misalignment* of electron lens elements, as well as of any constructional inaccuracies, is to spoil the axial symmetry of the system. Only two very simple instances of misalignment — a lateral shift and a tilt of a thin lens — will be considered qualitatively below.

A simple lateral shift Δ (Fig. 16·35) of the center of the lens with respect to the object signifies, on the one hand, that the center of the

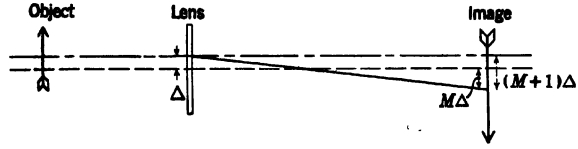


FIG. 16·35. Effect of Lateral Shift on Image Formation by a Thin Lens.

image is shifted by an amount $(|M| + 1)\Delta$, on the other, that this image center has the field aberrations characteristic of an object point a distance Δ off the axis. In the direction of the shift the range over which the image is satisfactory is thus reduced by the amount $(|M| + 1)\Delta$. The aberrations can, of course, be calculated by the usual formulas.

The tilt of an equipotential thin lens has a different effect. The center of the image is not shifted, but exhibits aberrations characteristic of an object point a distance $u\delta$ off the axis, u being the object distance and δ the angle of tilt. Furthermore, because of the relative inclination of the actual object and image planes to the conjugate planes of the tilted lens, off-axis points show an unsharpness even in the Gaussian approximation. Thus (Fig. 16·36) for object points a distance y_o from the axis in a meridional plane normal to the axis of tilt, the distance of the object points and image points from the conjugate planes is $y_o\delta$ and $My_o\delta$, respectively. Since $\Delta v = M^2\Delta u$, the radius of the consequent circle of confusion is

$$\Delta r = (M^2 y_o\delta + |M| y_o\delta)\frac{\theta}{M} = |M|\left(1 + \frac{1}{|M|}\right) y_o\delta\theta \qquad [16\cdot231]$$

where θ is the object-side aperture angle of the lens. In particular, the

[21] See section 14·3.

point corresponding to $y_o = \delta u$ exhibits only spherical aberration in addition to this defect, making the total aberration here

$$\Delta r = \left[|S_8| r_a^2 + |M| \left(1 + \frac{1}{|M|} \right) \delta^2 \right] r_a \qquad [16\cdot232]$$

Here $|M|$ is the absolute value of the magnification and r_a the aperture radius. Although the imaging in the plane containing the axis of tilt

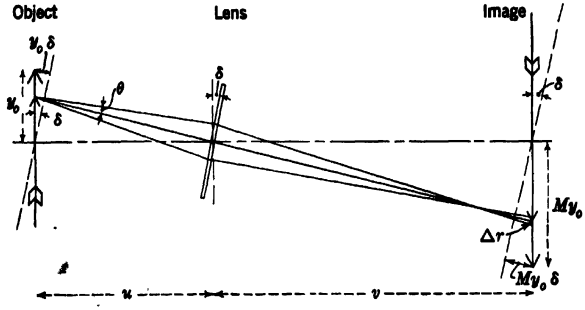


FIG. 16·36. Effect of Tilt on Image Formation by a Thin Lens.

also is astigmatic, it does not exhibit this added source of unsharpness and will show relatively little variation in sharpness — less in proportion as δ is larger.

REFERENCES

1. W. GLASER, "Theory of the electron microscope," *Z. Physik*, Vol. 83, pp. 104–122, 1933.
2. O. SCHERZER, "Some defects of electron lenses," *Z. Physik*, Vol. 101, pp. 593–603, 1936.
3. W. ROGOWSKI, "Aberrations of electron lenses," *Arch. Elektrotech.*, Vol. 31, pp. 555–593, 1937.
4. E. P. ADAMS, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, Smithsonian Institution, Washington, 1922.
5. K. STREHL, *Theorie des Fernrohrs auf Grund der Beugung des Lichts*, I. Theil, J. A. Barth, Leipzig, 1894.
6. O. SCHERZER, "Calculation of the third-order aberrations by the path method," in H. Busch and E. Brüche, *Beiträge zur Elektronenoptik*, J. A. Barth, Leipzig, pp. 33–41, 1937.
7. A. RECKNAGEL, "The theory of the electron mirror," *Z. Physik*, Vol. 104, pp. 381–394, 1937.
8. W. GLASER, "Chromatic aberration of electron lenses," *Z. Physik*, Vol. 116, pp. 56–67, June 1940.
9. E. G. RAMBERG, "Variations of the axial aberrations of electron lenses with lens strength," *J. Applied Phys.*, Vol. 13, pp. 582–594, September 1942.

10. V. K. Zworykin and G. A. Morton, "Applied electron optics," *J. Optical Soc. Am.*, Vol. 26, pp. 181–189, April 1936.
11. G. A. Morton and E. G. Ramberg, "Electron optics of an image tube," *Physics*, Vol. 7, pp. 451–459, December 1936.
12. J. B. Johnson, "A low-voltage cathode-ray oscillograph," *Phys. Rev.*, Vol. 17, pp. 420–421, 1921.
13. B. v. Borries and E. Ruska, "Short space-charge field of an auxiliary discharge as concentration lens for cathode rays," *Z. Physik*, Vol. 76, pp. 649–654, 1932.
14. E. E. Watson, "Dispersion of an electron beam," *Phil. Mag.* (7), Vol. 3, pp. 849–853, 1927.
15. B. v. Borries and J. Dosse, "Dispersion of electron rays due to their own space charge," *Arch. Elektrotech.*, Vol. 32, pp. 221–232, April 1938.
16. A. V. Haeff, "Space-charge effects in electron beams," *Proc. Inst. Radio Engrs.*, Vol. 27, pp. 586–602, September 1939.

CHAPTER 17

MAGNITUDE AND CORRECTION OF ELECTRON LENS DEFECTS

17·1. Approaches to the Aberrations Problem. The minimizing of lens defects or aberrations may well be regarded as the central problem of optical design. In electron optics workers have approached this problem from three different angles.

The first of these approaches consists in the determination of basic limits to the correction of electron-lens aberrations. It has already been mentioned that the existence of these limits, inherent in the properties of electric and magnetic fields in free space, necessitates the employment of relatively small apertures, hence justifying, in general, the restriction of attention to the aberrations of lowest order.

The second approach concerns itself with the determination, by analytical methods, of optimum systems under prescribed limiting conditions; for example, that the refracting field be short in extent compared with its focal length.

The last procedure consists in the measurement and calculation of the aberrations of practically useful systems, a comparison of the results suggesting lines of improvement or, at least, permitting the selection of the most favorable system examined. This last procedure comes closest to the process generally employed in the design of light-optical systems.

In this chapter the several aberrations are taken up individually and discussed from the three points of view, to the extent that they have become an object of study.

17·2. The Aperture Defect. The aperture defect or spherical aberration has received by far the greatest attention of all the aberrations. This is only reasonable; for in the two most important practical applications of electron optics — the electron gun and the electron microscope — it is the primary factor limiting the performance of the instrument. In the one case, the aperture defect of the " second lens " sets an upper limit to the current concentration attainable in a spot of given size; in the other, that of the objective determines the resolving power of the microscope.

The fact that the aperture defect of an electron lens forming a real image of an object (or a virtual image, provided that the imaging field is not cut off discontinuously at some point) cannot be made to vanish

was established by Scherzer.[1] Scherzer showed that the terms under the integral in the expression for the spherical aberration (Eq. 16·64*h*) could be arranged, after some further integrations by parts, as a sum of squares as follows:

$$\begin{aligned} S_8 = \frac{M}{16 r_{\alpha a}^3 \Phi_o^{1/2}} \int_{z_o}^{i} \Phi^{1/2} r_\alpha \Bigg\{ & \frac{5}{4}\left(\frac{\Phi''}{\Phi} + \frac{\Phi' r_\alpha'}{\Phi r_\alpha} - \frac{\Phi'^2}{\Phi^2}\right)^2 \\ & + \frac{\Phi'^2}{\Phi^2}\left(\frac{r_\alpha'}{r_\alpha} + \frac{7\Phi'}{8\Phi}\right)^2 \\ & + \frac{e}{mc^2\Phi}\left(H' + \frac{H r_\alpha'}{r_\alpha} - \frac{5H\Phi'}{4\Phi}\right)^2 + \frac{eH^2}{mc^2\Phi}\left(\frac{r_\alpha'}{r_\alpha} + \frac{\Phi'}{4\Phi}\right)^2 \\ & + \frac{\Phi'^4}{64\Phi^4} + \frac{e^2H^4}{4m^2c^4\Phi^2} + \frac{eH^2\Phi'^2}{32mc^2\Phi^3} \Bigg\} dz \end{aligned} \qquad [17\cdot1]$$

This expression can vanish only if both the electric and the magnetic fields are zero throughout the region between object and image; in short, if there is neither lens nor image.

Nevertheless, in theory at least, the aperture defect for given aperture angle may be reduced below any prescribed limit, even keeping the magnification and the separation of object and image constant.[2] If u is the object distance and θ the aperture angle of the imaging pencil, the spherical aberration may be written, making use of Eq. 17·1, if the aperture plane is thought of as coinciding with a principal plane of the lens system:

$$\Delta r_i = S_8 r_a^3 = u^3 S_8 \theta^3 = \frac{M u^3 r_{\alpha a} C \theta^3}{f^3} \qquad [17\cdot2]$$

Here C is a dimensionless constant characteristic of the lens, whose focal length is f, and of the position of the object (or image) relative to the lens. It is not affected by changes in the scale of the lens, provided that the relative dimensions as well as the electric and magnetic potentials of the electrodes and pole pieces are left unchanged.

Consider now two equipotential lenses (Fig. 17·1) of focal lengths f_1 and f_2, with the object placed at the focal point of the first, so that the image falls at the focal point of the second. The magnification of the compound system is $-f_2/f_1$. Furthermore, if C_1 and C_2 are the dimensionless characteristic constants of the two individual lenses with the object placed at their focal points, the aberration constant of the system

[1] See reference 1.

[2] See Rebsch, reference 2.

as a whole is

$$-u^3 S_8 = f_1^3 \cdot \frac{f_2}{f_1} \cdot f_1 \left(\frac{C_1}{f_1^3} + \frac{C_2}{f_2^3} \right) = f_2 C_1 + \frac{f_1^3}{f_2^2} C_2 \qquad [17\cdot3]$$

Let the two lenses be reduced to scale by a factor k, the two being pushed apart at the same time so as to leave the separation of object and image the same. This leaves the magnification unaltered, but causes the spherical aberration (Eq. 17·3) to be reduced by the same factor k. Since k is arbitrary, the spherical aberration of the system may be made arbitrarily small in this manner.

In actual practice the reduction of spherical aberration by decreasing the scale of the electron lens cannot be pushed very far. A reduction in dimensions by the factor k would require an increase in the strength of

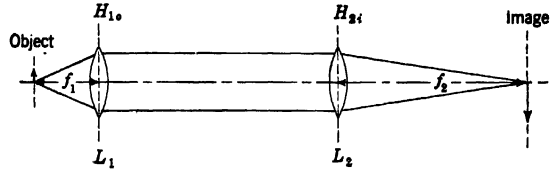


FIG. 17·1. Reduction of Spherical Aberration below Any Prescribed Amount.

the magnetic or electric lens fields by a factor $1/k$. Such an increase is limited, however, on the one hand by the tendency of ferromagnetic materials to become saturated at high fields, on the other, by cold emission from the electrodes of the lens.

For the sake of completeness, it should be mentioned that magnetic and electric fields free of spherical aberration exist — at least in principle — and have been derived by Glaser[3] and Recknagel.[4] For example, by suitable integrations by parts the expression for the spherical aberration in Eq. 16·60h can be converted, for a purely magnetic field, into the form

$$S_8 = \frac{M}{96 r_{ca}^3 \Phi} \int_{z_o}^{z_i} r_\alpha^4 \left\{ \frac{2e^2 H^4}{m^2 c^4 \Phi} + \frac{5eH'^2}{mc^2} - \frac{eHH''}{mc^2} \right\} dz \qquad [17\cdot4]$$

Setting the integrand equal to zero leads to a differential equation for the magnetic field, which is solved by a family of functions of the type plotted in Fig. 17·2. These fields are finite in extent, their length being related to their strength in such fashion that they are in no case capable of reducing the initial slope of the rays leaving the object (both object

[3] See reference 3.
[4] See reference 4.

and aperture plane must lie within the field, the field being cut off sharply at the aperture plane) by more than 5 per cent. Accordingly, the use of these fields does not offer much promise for the correction of spherical aberration in real images. Similar conditions hold for the aberration-free electrostatic fields.

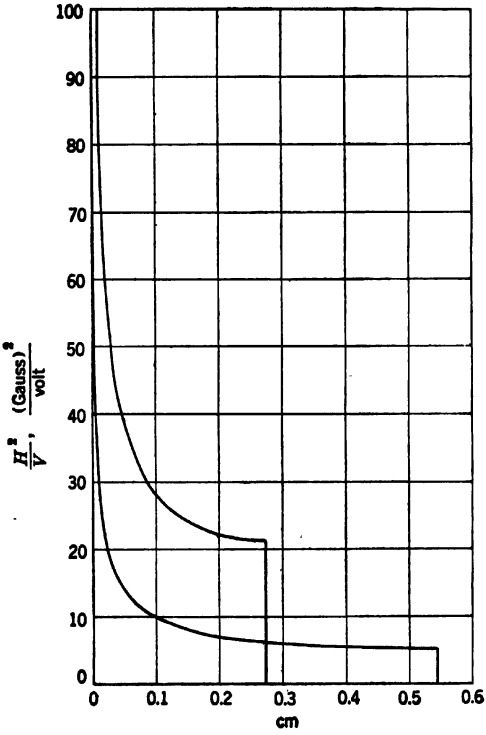


FIG. 17·2. Magnetic Fields Free from Spherical Aberration. (Glaser, reference 3.)

The problem resolves itself, thus, into a search of systems with a minimum of spherical aberration rather than of aberration-free lenses. At the same time, the fact that the spherical aberration can be made arbitrarily small by reducing the dimensions of the imaging system to scale indicates that restrictions must be placed on any set of systems which are to be compared. An inspection of the coefficient S_3 in Eq. 17·1 shows that both increasing the focal length (decreasing Φ'^2/Φ^2 and H^2/Φ^2 by a constant factor) and increasing the lens thickness without changing the focal length (multiplying z by $k > 1$ and both Φ'^2/Φ^2 and

H^2/Φ by $1/k$) cause a reduction in S_8, provided that r_a does not vary too rapidly within the lens. When, in particular, different weak and thin lenses are compared, it is thus logical to require that the focal lengths and the thicknesses of the systems to be compared be identical.

By writing, in accord with Eq. 15·62,

$$T = \frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 + \frac{eH^2}{8mc^2\Phi}$$

a suitable definition for the thickness l of the lens is

$$\left(\frac{l}{2}\right)^2 = \frac{\int_{-\infty}^{\infty} T \cdot z^2 \cdot dz}{\int_{-\infty}^{\infty} T \cdot dz} \qquad [17·5]$$

The zero plane of z has here been chosen so that

$$\int_{-\infty}^{\infty} T \cdot z \cdot dz = 0$$

that is, through the " center of gravity " of the function $T(z)$.

Consider now only short and weak electron lenses, whose axial potentials and magnetic field strengths may be expressed by

$$\Phi = \Phi_A[1 + A\phi(z)] \qquad |A\phi| \ll 1 \qquad [17·6a]$$

$$H = H_{\max} \cdot \frac{H}{H_{\max}} \qquad \frac{el^2H_{\max}^2}{8mc^2\Phi_A} \ll 1 \qquad [17·6b]$$

With only the lowest powers of A and $H_{\max}/\Phi_A^{1/2}$ retained, the refractive power $1/f$, the thickness l, and the spherical-aberration coefficient are given by the following expressions:

$$\frac{1}{f} = \frac{3A^2}{16}\int_{z_o}^{z_i} \phi'^2\, dz + \frac{eH_{\max}^2}{8mc^2\Phi_A}\int_{z_o}^{z_i} \frac{H^2}{H_{\max}^2}\, dz \qquad [17·7]$$

$$\frac{1}{f}\left(\frac{l}{2}\right)^2 = \frac{3A^2}{16}\int_{z_o}^{z_i} \phi'^2 z^2\, dz + \frac{H_{\max}^2}{8mc^2\Phi_A}\int_{z_o}^{z_i} \frac{H^2}{H_{\max}^2}\, z^2\, dz \qquad [17·8]$$

$$S_8 = \frac{5bA^2}{64}\int_{z_o}^{z_i} \phi''^2\, dz + b\,\frac{H_{\max}^2}{8mc^2\Phi_A}\int_{z_o}^{z_i} \frac{H'^2}{H_{\max}^2}\, dz \qquad [17·9]$$

where b is the image distance.

From these expressions it is seen that the ratio

$$G = \frac{b}{S_8 f l^2} \qquad [17·10]$$

is dimensionless and independent of the absolute magnitude of A and $H_{max}/\Phi_A^{1/2}$ (though dependent on their ratio) as well as of the magnification of the image. It may thus properly be regarded, as suggested by Rebsch and Schneider,[5] as a figure of merit of any weak, thin lens with regard to the aperture defect.

The problem of minimizing the spherical-aberration coefficient, Eq. 17·9, keeping the focal length and thickness of the systems to be compared constant, was solved for weak, thin lenses by Scherzer[6] and Rebsch and Schneider.[5] Their findings may be summarized as follows:

a. Optimum electric unipotential lens, $\phi = e^{-Bz^2}$.	$G = 0.267$
b. Optimum electric immersion lens, $\phi' = e^{-Bz^2}$.	$G = 2.4$
c. Optimum magnetic lens, $H = H_{max}e^{-Bz^2}$	$G = 2.0$
d. Optimum rotation-free magnetic lens, $H = \text{const}\ ze^{-Bz^2}$.	$G = 0.222$

For similar variation of the field (ϕ' and H/H_{max}, respectively) the figure of merit of the weak magnetic lenses is thus five-sixths that of the corresponding weak electric lenses — rotation-free magnetic lenses corresponding mathematically to unipotential electric lenses.

Rebsch and Schneider showed furthermore that the figure of merit of a compound electric and magnetic lens is always less than the larger of the two figures of merit, G_{el} and G_{magn}, of the electric- and the magnetic-lens component by itself. The figure of merit of the compound lens is given by

$$G_{comp} = G_{el}\frac{\left(1+\eta^2\frac{f_{el}}{f_{magn}}\right)^2}{\left(1+\eta^2\frac{S_{8magn}}{S_{8el}}\right)\left(1+\eta^2\left(\frac{f_{el}}{f_{magn}}\right)\left(\frac{l_{magn}}{l_{el}}\right)^2\right)} \qquad [17\cdot11]$$

where f_{el}, f_{magn}, S_{8el}, S_{8magn}, l_{el}, l_{magn} are calculated for $A = (eH_{max}^2/[mc^2\Phi_A])^{1/2} = 1$ and $\eta = (eH_{max}^2/[mc^2\Phi_A])^{1/2}/A$ measures the relative strength of the magnetic and electric fields.

The weak-lens figures of merit of a number of other lens fields are given below, to illustrate the closeness with which commonly employed configurations approach the optimum. The first four are taken from the paper of Rebsch and Schneider.

e. Charged ring of radius a, $\phi = 1/(a^2+z^2)^{1/2}$.	$G = 0.119$
f. Charged ring of radius a inserted in aperture of nearly equal radius, $\phi = 1/(a^2+z^2)$.	$G = 0.200$

[5] See reference 5.
[6] See reference 6.

g. Two oppositely charged rings of equal radius placed close together, $\phi = z/(a^2 + z^2)^{3/2}$. $G = 0.239$

h. $\phi = 1/(a^2 + z^2)^n$.

$$G = \{2(2n + 3)(4n - 1)\}/\{15(4n + 1)(n + 1)\}$$

i. $\phi = A\{e^{-(Bz+0.85)^2/2} + e^{-(Bz-0.85)^2/2}\}$. $G = 0.226$

j. Immersion lens consisting of adjoining equidiameter cylinders of radius a, $\phi \cong \tanh(1.315z/a)$. $G = 2.33$

k. Magnetic lens consisting of adjoining cylindrical pole pieces of equal internal diameter (radius a), $H \cong H_{\max} \operatorname{sech}^2(1.315z/a)$. $G = 1.94$

l. Magnetic lens formed by single circular current loop of radius a, $H = H_{\max}\, a^3/(a^2 + z^2)^{3/2}$. $G = 1.6$

The two-cylinder immersion lens and its magnetic counterpart, in particular, are seen to approach the optimum values of the figure of merit very closely.

It should be mentioned that the relative magnitudes of the figures of merit depend materially on the definition of lens thickness. For example, if, as has been suggested by Plass,[7] the thickness of symmetrical electric lenses is defined as the distance between the two points at which ϕ differs from $\phi(\infty)$ by one one-hundredth of the difference between the value of ϕ at the center of symmetry and $\phi(\infty)$, the figure of merit of the system *i* becomes actually about 22 per cent greater than that of the "optimum" system *a*.

Apart from this, it must be remembered that the electron lenses employed most widely are not weak lenses but strong lenses. Hence, the weak-lens calculations are important only to the extent in which their results can be transferred to the case of strong lenses.

Consider, in particular, lenses in which the object is placed near the focal point, that is, where large magnification is to be achieved, as in the electron microscope. Here $b = Mf$, so that

$$G = \frac{M}{S_8 l^2} \qquad [17\cdot12]$$

Since l^2, in accordance with its definition by Eq. 17·5, is independent of the strength of the lens (for weak lenses), the more readily interpreted quantity $S_8/M \propto 1/G$, indicating directly the size of the circle of confusion, due to spherical aberration, referred back to the object, is conveniently used as a measure of the quality of the lens at any given lens strength.

Figure 17·3[8] shows four electron lenses — two unipotential lenses, an

[7] See reference 7.

[8] See Ramberg, reference 8.

immersion lens, and a magnetic lens — for which the variation of the spherical-aberration coefficient S_3/M with lens strength (as given by the voltages applied to the electrodes and the ratio H^2_{max}/V, respectively)

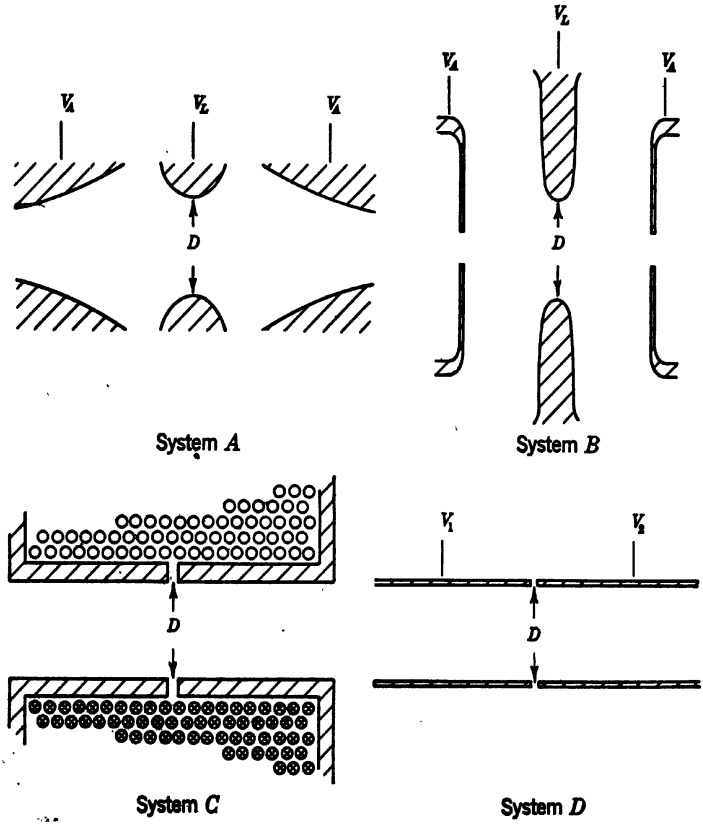


FIG. 17·3. Typical Electron Lenses. (Courtesy of *J. Applied Phys.*, reference 8.)

was calculated by integrating Eq. 15·62 and evaluating the integral in Eq. 17·1. System *A*, with the potential distribution

$$\frac{\Phi}{V_A} = 1 - 0.827\left(\frac{V_A - V_L}{V_A}\right)e^{-1.445s^2} \qquad [17·13]$$

the diameter of the central electrode being used as unit of length as in the remaining three systems, is Scherzer's "unipotential lens of least

spherical aberration." The electrode configuration shown was determined by Plass[9] from Eq. 11·75.

System B is the unipotential lens whose field distribution is shown in Fig. 13·10a. Its weak-lens figure of merit is $G = 0.194$.

System C is the magnetic lens with the axial field distribution

$$H \cong H_{\max} \operatorname{sech}^2(2.630z) \qquad [17\cdot14]$$

Finally, the field distribution for the two-equidiameter-cylinder immersion lens System D is given by

$$\frac{2\Phi}{V_2 + V_1} \cong 1 + \left(\frac{V_2 - V_1}{V_2 + V_1}\right) \tanh\,(2.630z) \qquad [17\cdot15]$$

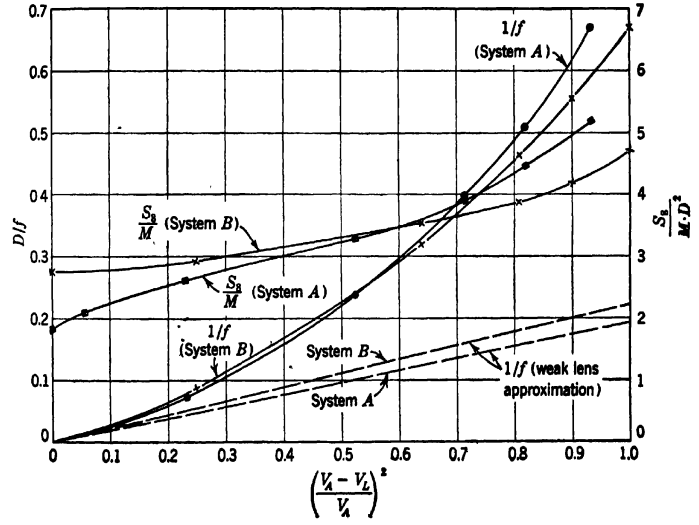


FIG. 17·4. Variation of Refractive Power and Spherical Aberration of Unipotential Electric Lenses with Electrode Voltages. (Courtesy of *J. Applied Phys.*, reference 8.)

The results for the two unipotential lenses, Systems A and B, are shown in Fig. 17·4. In both cases it is assumed that the center electrode is negative with respect to the outer electrodes. A plot of the refractive power is included. Both the refractive powers and the spherical-aberration coefficients S_8/M are seen to be larger throughout than their weak-lens approximations, the deviations of the refractive power exceeding those of S_8/M. The most interesting feature of the graph is that the spherical-aberration coefficient for System B, although initially larger than that for System A, becomes smaller than the latter at high

[9] See reference 7.

lens strengths. Similarly, System i on p. 609, with $B = 2.53$ to yield the same (weak-lens) thickness, l, as that of System A, becomes superior to System A at high lens strengths. At $1/f = 0.685$ diameter^{-1}, S_8/M for System i is 3.72 diameter^{-2}, whereas for System A, $S_8/M = 5.25$ diameter^{-2}. It appears that a less sharply peaked potential distribution is more favorable at the high lens strengths.

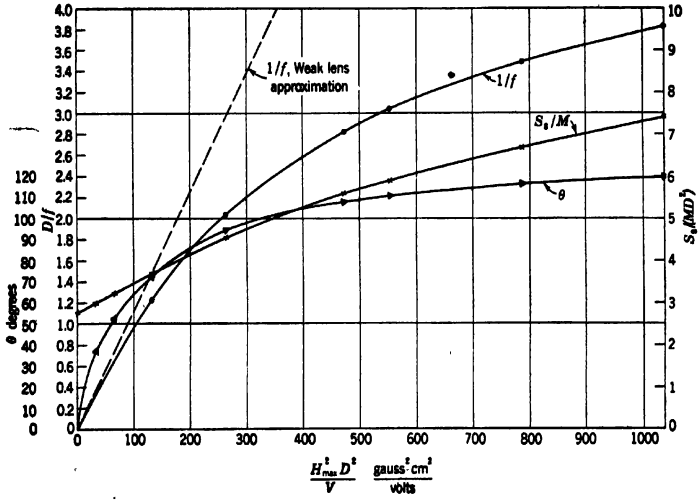


FIG. 17·5. Variation of Refractive Power and Spherical Aberration of Magnetic Lens (System C) with Maximum Field Strength. (Courtesy of *J. Applied Phys.*, reference 8.)

Figure 17·5 shows corresponding curves for the magnetic lens, System C. Here the refractive power, instead of exceeding the weak-lens value, remains consistently below it. The curve of the spherical-aberration coefficient, similarly, is convex to the axis of the abscissas. These effects must in part be ascribed to the fact that a progressively larger part of the magnetic field ceases to contribute to the lens action as the object plane moves into the lens field. A curve indicating the rotation of the image, θ, is included.

Figure 17·6, finally, reproduces the results for the immersion lens, System D. Here the weak-lens approximations to the focal lengths are

$$\frac{1}{f_1} = 0.658\left(\frac{\Phi_2}{\Phi_1}\right)^{1/4}\left(\frac{V_2 - V_1}{V_2 + V_1}\right)^2 \text{diam}^{-1}$$
$$\frac{1}{f_2} = 0.658\left(\frac{\Phi_1}{\Phi_2}\right)^{1/4}\left(\frac{V_2 - V_1}{V_2 + V_1}\right)^2 \text{diam}^{-1} \qquad [17\cdot16]$$

The actual refractive powers exceed these values throughout. However, the spherical-aberration coefficients S_{s1}/M and S_{s2}/M — the subscript 1 indicating an accelerating, the subscript 2 a decelerating

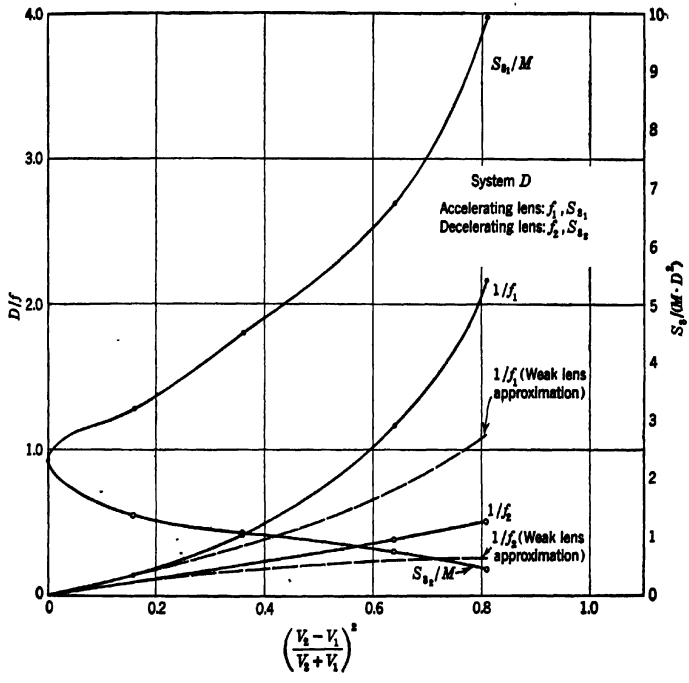


FIG. 17·6. Variation of Refractive Power and Spherical Aberration of Immersion Lens (System *D*) with Applied Voltages. (Courtesy of *J. Applied Phys.*, reference 8.)

lens — increase and decrease, respectively, with increasing lens strength. At the same time the quantities $(S_{s1}/M)(\Phi_1/\Phi_2)^{1/4}$ and $(S_{s2}/M)(\Phi_2/\Phi_1)^{1/4}$ are found to differ less than 10 per cent from their weak-lens value throughout the range covered.

In electron-microscope work the spherical-aberration coefficient determining the limit of resolution obtainable with a given objective is not S_s/M, but $f^3 S_s/M = Cf$ (C being defined by Eq. 17·2). This is the ratio of the aberration Δr to the cube of the aperture angle. Accordingly, Fig. 17·7 replots the results for the spherical aberration of the four systems *A* to *D*, showing Cf as a function of the refractive power $1/f$. The scale $d_{\min}$ indicates relative values of the resolution which may be

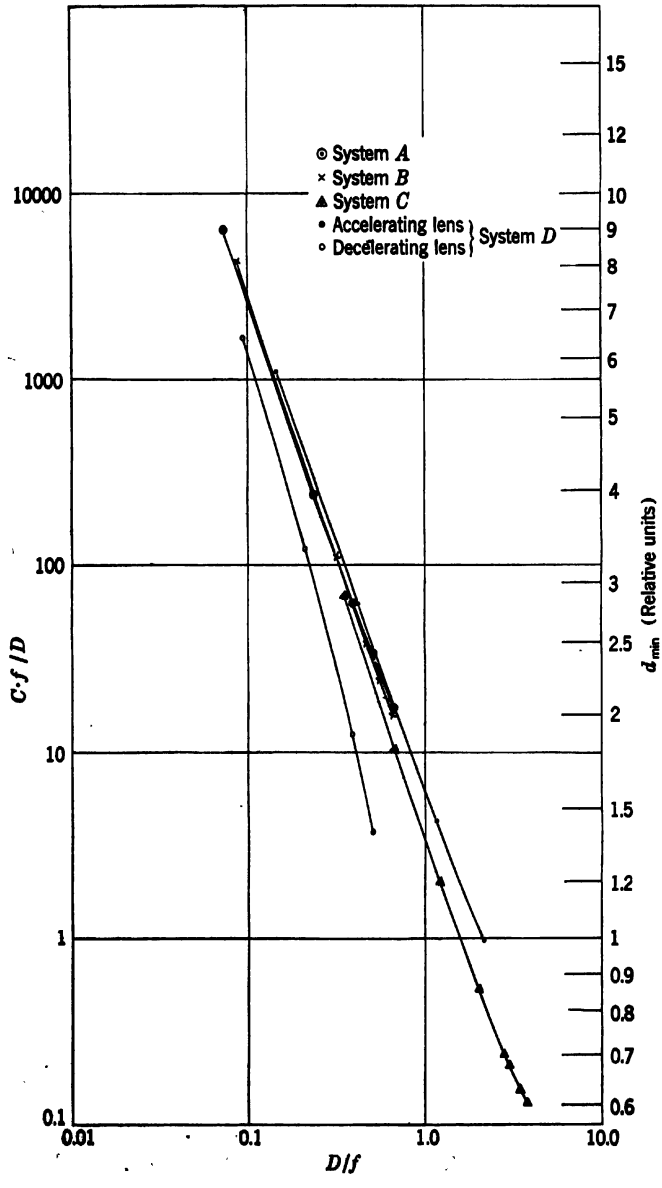


FIG. 17·7. Variation of Spherical Aberration Coefficient Cf and of Resolving Power with Refractive Power. (Courtesy of *J. Applied Phys.*, reference 8.)

attained, it being assumed that the resolution is proportional to $(Cf)^{\frac{1}{4}}$. Cf is seen to decrease (for the equipotential systems) somewhat more slowly than $(1/f)^{-3}$, more nearly as $(1/f)^{-2\frac{1}{2}}$, over practically the entire range covered.

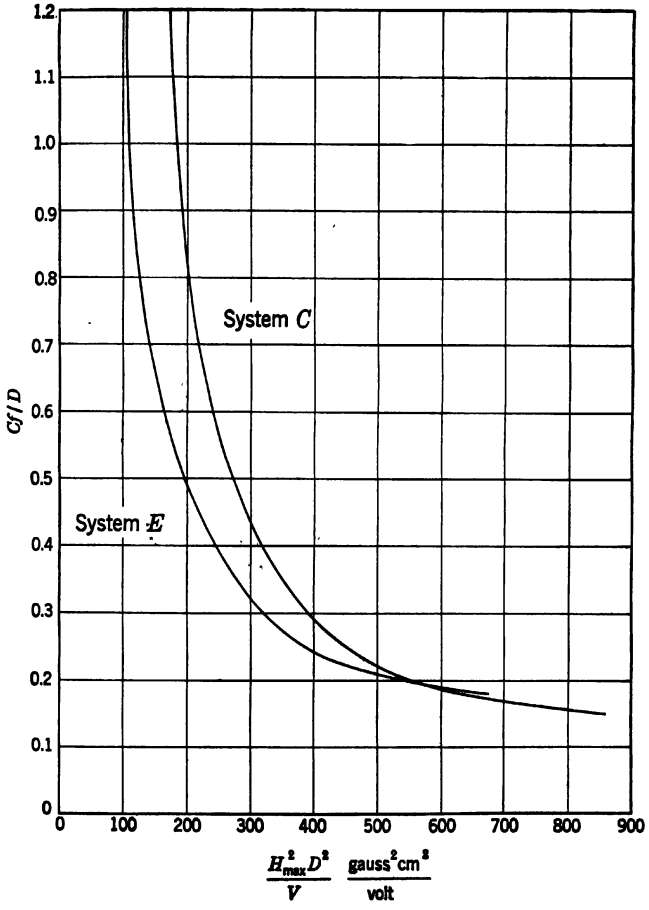


FIG. 17·8. Variation of Spherical Aberration with Maximum Axial Field Strength for Two Magnetic Lenses.

Figure 17·8 compares the spherical aberration of an actual asymmetric magnetic electron-microscope objective (System E) with that of System C. The calculations for System E, shown in Fig. 4·13 (in the range

below saturation), were carried out by Dosse[10] on the basis of measured field values. The results for the two systems are seen to be quite comparable. For System E the diameter D of the smaller pole piece serves as unit of length.

In a number of applications of electron lenses other than the electron microscope — in particular in the electron gun — lenses are frequently employed which are short, that is, consist of refractive fields much shorter than the focal length of the lens, and yet cannot be regarded as weak. Magnetic lenses of this type have been studied by Glaser.[11] Equation 17·4 may be written

$$S_8 = \frac{eM}{16 r_{aa}^3 mc^2 \Phi} \int_{z_o}^{z_i} \left\{ H'^2 + \frac{eH^4}{3mc^2\Phi} - \frac{(H^2)''}{12} \right\} r_a^4 \, dz \qquad [17\cdot17]$$

If, for a short lens, r_a^4 is treated as a constant ($= r_{aa}^4$), this becomes

$$S_8 = -\frac{eb}{16mc^2\Phi} \int_{-\infty}^{\infty} \left\{ H'^2 + \frac{eH^4}{3mc^2\Phi} \right\} dz \qquad [17\cdot18]$$

the integral of the last term vanishing at the limits of integration; b is the image distance. Finally, if the expression for the refractive power $1/f$ of a short magnetic lens (Eq. 15·86) is utilized, the dimension-free aberration constant C defined by Eq. 17·2 is given by

$$C = \frac{S_8 f^3}{b} = -\left(\frac{J_o J_2}{128}\right) \kappa^2 - \left(\frac{J_4}{6J_2}\right) \kappa$$

$$J_o = \int_{-\infty}^{\infty} \frac{H'^2}{H_{\max}^2} dz \qquad [17\cdot19]$$

$$J_2 = \int_{-\infty}^{\infty} \frac{H^2}{H_{\max}^2} dz \qquad J_4 = \int_{-\infty}^{\infty} \frac{H^4}{H_{\max}^4} dz$$

where the parameter κ is given by

$$\kappa = \frac{eH_{\max}^2 f^2}{mc^2\Phi} \qquad [17\cdot20]$$

κ is proportional to the focal length of the lens. The relation Eq. 17·19 is in accord with the previously mentioned finding that for a number of (not necessarily short) systems Cf is proportional to $f^{4/3}$, that is, to a power of the focal length between the second and the third.

The aberration coefficients C for a number of short magnetic lens fields, as determined by Glaser, are given below. The formulas are valid for $\kappa > 20$ corresponding to a ratio of the focal length to the thick-

[10] See reference 9.

[11] See reference 10.

ness $f/l \geq 2$ (in the case of the exponential field, System q, more exactly $f/l \geq 1.8$). Here l is the thickness as defined by Eq. 17·5. The value of C for $\kappa = 20$ is added in parentheses:

$$m.\quad \frac{H}{H_{\max}} = 1 - \left(\frac{z}{a}\right)^2 \qquad |z| < a \qquad C = -0.299\kappa - 0.0222\kappa^2 \quad (14.9)$$

$$= 0 \qquad |z| > a$$

$$n.\quad \frac{H}{H_{\max}} = \cos\left(\frac{\pi z}{|2a|}\right) \qquad |z| < a \qquad C = -0.125\kappa - 0.0177\kappa^2 \quad (9.6)$$

$$= 0 \qquad |z| > a$$

$$o.\quad \frac{H}{H_{\max}} = e^{-(z/a)^2} \qquad C = -0.117\kappa - 0.0122\kappa^2 \quad (7.2)$$

$$p.\quad \text{Circular current loop: } \frac{H}{H_{\max}} = (1 + [z/a]^2)^{-3/2} \qquad C = -0.109\kappa - 0.0101\kappa^2 \quad (6.2)$$

$$q.\quad \frac{H}{H_{\max}} = e^{-|z/a|} \qquad C = -0.083\kappa - 0.007\kappa^2 \quad (4.8)$$

Glaser showed that a field closely approximated by System q, having a sharp peak at the center, represents the best short-lens field from the point of view of spherical aberration. It may be mentioned that System q also has the same weak-lens figure of merit, $G = 2.0$, as the optimum (analytic) weak-lens field (System o) according to Rebsch and Schneider's criterion.

Although analytical methods are valuable in establishing limiting values of the aberrations, deducing general relationships and indicating favorable conditions for low lens aberration, the determination of the aperture defect and other aberrations for actual lenses must, more often than not, be accomplished partly or wholly experimentally. This is especially true of strong iron-encased magnetic lenses, in which the complex properties of ferromagnetic materials make field calculations difficult and the small dimensions of the pole pieces render measurements of the field distribution along the axis inaccurate.

In the case of electric lenses, almost all the measurements of spherical aberration have been carried out for systems consisting of two coaxial cylinders at different potentials. Epstein[12] and Gundert[13] determined the aberration by inserting a diaphragm containing a set of small circular apertures placed along two radii from the center outward in one of the

[12] See reference 11.

[13] See references 12 and 13.

cylinders and by measuring the distance of the luminous spots due to each aperture from the optic axis in the plane of paraxial focus. Figure 17·9 shows the arrangement employed by Epstein. A somewhat different method, which is especially useful for the determination of the spherical aberration of short-focus lenses, has been employed by Spangenberg and Field.[14]

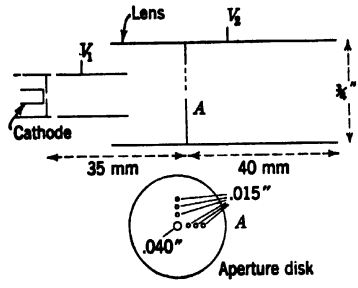


FIG. 17·9. Arrangement for Measuring Spherical Aberration of Two-Cylinder Immersion Lens. (Epstein, reference 11.)

Consider, in Fig. 17·10, an electron lens, whose position is indicated by its "principal planes," H_i and H_o, and, on either side, a regular fine-meshed screen, S_1 and S_2, respectively. Place a point source P of electrons at the image point (or at the object point, if the latter is further removed from the lens than the image point), a distance b from H_i and a distance s_1 from S_1. Furthermore, place a fluorescent screen (or photographic plate) a distance L from H_o and a distance s_2 from S_2. Let the height of incidence of a given ray from the source P on the "principal plane" H_i (and hence also on H_o) be r_a. Then $r_a = r_1 \cdot b/s_1$, where r_1 is the height of incidence on the screen S_1. This assumes that the principal plane H_i is, in fact, a plane.

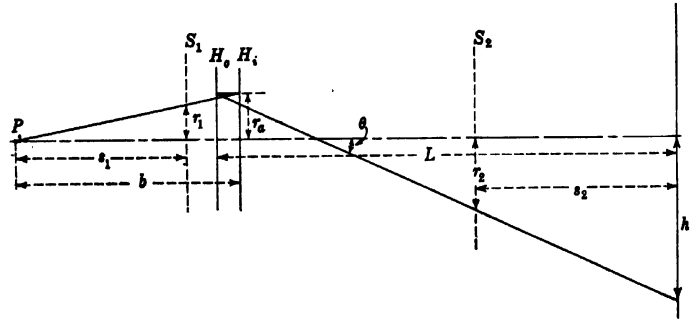


FIG. 17·10. Experimental Determination of Spherical Aberration. (Spangenberg and Field, reference 14.)

Such an assumption is not generally valid in the calculation of aberrations. However, if the distance b is large, it leads to a negligible error. It will not be necessary to make a similar assumption for the "principal

[14] See reference 14.

plane" H_o. The value of r_1, and hence the value of r_a, can thus be determined by counting the meshes on the image of S_1 on the screen between the axis and the point, a distance h from the axis, at which the ray traverses the fluorescent screen. The distance from the axis at which the ray passes through the screen S_2, r_2, is similarly determined from its projection image on the fluorescent screen or photographic plate. From it the inclination of the ray in the region between the lens and the observation screen, $\theta = (h - r_2)/s_2$, is found. Finally,

$$\frac{h}{\theta} = \frac{hs_2}{h - r_2} = L - u - \Delta u = L - u + \frac{S_8 r_a^2 u}{M}$$

$$= L - u + \left(\frac{S_8}{M}\right) u^3\theta^2 \qquad [17\cdot21]$$

establishes the position of the point of convergence for any given angle θ. By plotting h/θ against θ, a parabolic curve is obtained, whose vertex indicates the position of the paraxial focus relative to the observation screen, $L - u$. The deviations from this limiting value determine the magnitude of the spherical-aberration coefficient S_8/M. The object distance u is given by

$$\lim_{\theta \to 0}\left(\frac{r_a}{\theta}\right) = \lim_{h \to 0}\left(\frac{r_1 \cdot b \cdot s_2}{s_1(h - r_2)}\right)$$

In particular, in the case of the incidence of a parallel beam of electrons on the lens and on screen S_1, a knowledge of the relative positions of the screen S_2, the observation screen and some reference plane attached to the lens, together with a single photograph of the projected screen patterns suffice to determine exactly the position of the focal point, the focal length of the lens, and the spherical-aberration coefficient of the lens for large magnifications. Once the focal length of the lens, f, has been determined, the spherical-aberration constant is obtained from the distortion of the pattern of screen S_2 alone. From Fig. 17·11 it follows that

$$\frac{h - r_2}{s_2} = \frac{h}{L - f + Cf\theta^2} \qquad [17\cdot22]$$

$$C = \frac{L - f}{f}\frac{s_2^2}{(h - r_2)^2}\left(\frac{s_2}{L - f}\frac{h}{h - r_2} - 1\right) \qquad [17\cdot23]$$

It should be noted that the distortion of the shadow image of S_1 becomes a measure of the spherical aberration only if the principal plane H_o is,

indeed, a plane, a condition normally not fulfilled. Under these circumstances

$$-\frac{h}{L-f+Cf\theta^2}=\frac{r_1}{f-Cf\theta^2} \qquad [17\cdot24]$$

$$C=\frac{L^2r_1}{(h-r_1)^3}\left(\frac{h}{r_1}+\frac{L-f}{f}\right) \qquad [17\cdot25]$$

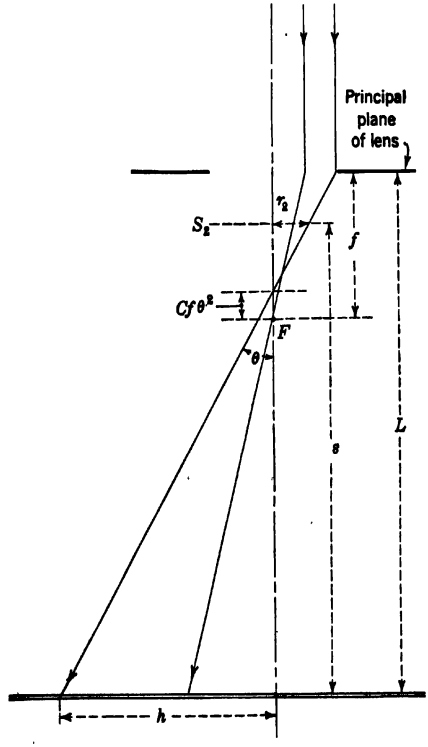


FIG. 17·11. Determination of Spherical Aberration with Parallel Incident Pencil.

To return to the generally valid Eq. 17·23, it may be converted into the form

$$C\cong\frac{1}{M}\cdot\frac{\Delta M}{M}\frac{(L-f)^4}{fs_2h^2} \qquad [17\cdot26]$$

if the magnification M of the pattern of S_2 and the difference in magnification ΔM of an object on S_2 at the distance r_2 from the axis and one on

the axis is introduced, the magnification being measured at right angles to the radius vector r_2. Dosse[15] employed a spider thread 0.9 micron in diameter as object in determining the spherical aberration of a microscope objective. This was passed through the objective with the aid of a special object holder with a long tubular prolongation. If the thickness of the thread at the center of the image is d_o, that at a distance h from the center, d_1, $\Delta M/M = (d_1 - d_o)/d_o$. Since it was not possible to bring the object into the field-free region beyond the lens, a small correction had to be applied for the curvature of the rays between the object and the fluorescent screen. The value of the spherical-aberration coefficient of an objective (System E) with a focal length of 2.7 millimeters was thus found to be $C = 0.41$. Dosse's calculations had yielded $C = 0.59$. In view of the errors involved both in the measurement and in the calculation, the agreement may be regarded as adequate.

To return to the measurements of spherical aberration on two-cylinder electrostatic immersion lenses, the findings of Epstein and Spangenberg and Field for a diameter ratio $D_2/D_1 = 1.5$ agree within experimental error with the theoretical results for $D_2/D_1 = 1$ shown in Fig. 17·6. The more extensive measurements of Gundert indicate that, with the diameter of the high-voltage cylinder held constant, the least aberration results if both cylinders are of the same diameter. Furthermore, as might be expected, an increase in the diameter of the low-voltage cylinder has a smaller adverse effect on the aberration than a reduction in diameter. A lens formed by a pair of apertures was found to have practically the same aberration as one formed by a pair of cylinders of equal diameter.

A measurement of spherical aberration on an electrostatic lens of different type, that is, a short-focus unipotential lens such as that shown in Fig. 13·10*b*, has been carried out by von Ardenne.[16] The diameter of the central aperture was 0.125 millimeter and the focal length 3 millimeters. The measurement was made by examining the image of a pointlike distant electron source formed by the lens on uranium glass with the aid of a light microscope. In this manner von Ardenne found the value 37 for the dimension-free spherical-aberration constant C (Eq. 17·2). A calculation for a similar system by one of the authors yielded the larger value $C = 179$.

Measurements of the aperture defect of magnetic lenses have been reported by E. Ruska[17] and Becker and Wallraff.[18] Ruska's measure-

[15] See reference 9.
[16] See reference 15.
[17] See reference 16.
[18] See reference 17.

ments were made on a typical objective lens (Fig. 4·7) with variable pole-piece separations. He observed the image of a fine aperture illuminated by a diffuse electron beam formed by the lens provided with a suitable ring aperture on a luminescent screen and read the coil current values I_r required to focus the image for ring apertures of different radius r. Then, if the pole pieces are assumed to be unsaturated, the ratio of the focal lengths for the zone of radius r and the central zone is given, for relatively long focal lengths, by

$$\frac{f_r}{f_o} = \frac{I_r^2}{I_o^2}$$

From this the spherical aberration constant S_8/M may be determined from the approximately correct relation

$$\frac{f_r}{f_o} = 1 - \frac{S_8 r^2}{M}$$

The measurements indicated that the spherical aberration is reduced as the pole-piece separation is increased. Thus, by using again the least internal diameter of the pole pieces as unit of length, it was found that the pole-piece support alone, with $1/f \cong 2$ diameters^{-1}, gave $S_8/M = 5$ diameters^{-2}. The pole pieces in conventional position, that is, with the conic surfaces outward, gave, for $1/f = 0.7$ diameter^{-1}, $S_8/M = 3.9$ for a pole-piece separation of 0.29 diameter and $S_8/M = 3.0$ for a pole-piece separation of 0.86 diameter. For the same approximate value of the focal length ($1/f = 0.7$ diameter^{-1}) the pole pieces, placed so that their tips (and flat surfaces) were turned outward and a distance of 2.2 diameters apart, gave the much lower value $S_8/M = 1.1$ diameters^{-2}. All the results except the last agree well with the value of the spherical aberration calculated for the idealized two-cylinder magnetic lens (Fig. 17·5). Ruska's measurements were carried out with accelerating voltages between 40 and 60 kilovolts.

Becker and Wallraff's measurements were made at lower voltages — 4 to 20 kilovolts — on a series of coils with an inner diameter between 80 and 120 millimeters, one of which was encased in an iron shield with a variable gap. Their method of measurement resembled Ruska's, with the exception that they used a movable screen to determine the change in focus with the effective zone radius of the lens, thus freeing themselves from the major sources of error of Ruska's method.

For iron-free wire coils of the same diameter and focal length they found the interesting relation

$$S_8 = \text{const } d^{-0.4}$$

where d is the length, measured along the axis, of the coil. Thus, the spherical aberration is found to decrease with increasing coil length somewhat more slowly than in proportion to the inverse square root of the coil length. The coil length was varied between 20 and 600 millimeters.

Figure 17·12 shows the variation of the longitudinal spherical aberration of the iron-enclosed coil with the gap width. This coil has an inner diameter of 80 millimeters and a length of 65 millimeters. Measurements at very small gap widths suggested that the aberration is least for very narrow gaps. For example, for $r_a = 11$ millimeters, the aberration for a 1-millimeter gap was found to be equal to that for a 55-millimeter gap. At the same time the lens current required to obtain a given focal length, which attains a minimum at a gap width of about 40 millimeters, is enormously greater at the small gap width. Hence gap widths only slightly smaller than the inner pole-piece diameter are generally most desirable. For very strong lenses with very small clear diameters saturation phenomena modify the dependence of the aberration on the gap width.

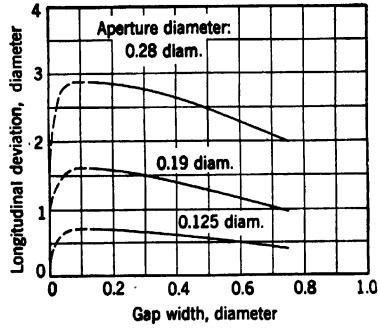


FIG. 17·12. Variation of Longitudinal Spherical Aberration with Width of Gap of Iron-Enclosed Coil. (Becker and Wallraff, reference 17.)

For a few very special electric and magnetic fields the spherical aberration can be determined analytically. In particular, the uniform magnetic and electric fields belong to this class. Thus it has already been shown (Eq. 15·33) that the spherical aberration of a uniform magnetic field is given by

$$\Delta r = -\frac{\pi c}{H}\left(\frac{2m\Phi}{e}\right)^{1/2}\theta_o^3 = -\frac{10.54\Phi^{1/2}}{H}\theta_o^3 \text{ cm} \qquad [17\cdot27]$$

where θ_o is the angle of inclination of the ray at the object. Similarly, the spherical aberration of a superposed uniform electric and magnetic field is, according to Eq. 15·170,

$$\Delta r = -\frac{10.54\Phi_o^{1/2}}{H}\left(\frac{\Phi_o}{\Phi_i}\right)^{1/2}\theta_o^3 \text{ cm} \qquad [17\cdot28]$$

To these may be added the formula for the simple uniform electric

field of length L, which may be derived in similarly elementary fashion:

$$\Delta r = \frac{L\Phi_o}{\Phi_i - \Phi_o}\left(1 - 2\left[\frac{\Phi_o}{\Phi_i}\right]^{\frac{1}{2}} + \frac{\Phi_o}{\Phi_i}\right)\theta_o^3 \qquad [17\cdot29]$$

In this last case the image is virtual and, hence, the sign of the spherical aberration positive instead of negative. All the three fields above form images of unity magnification.

A more complex field for which the spherical aberration has been determined analytically is the " bell-shaped " field

$$H = \frac{H_o}{1 + \left(\frac{z}{a}\right)^2} \qquad [17\cdot30]$$

studied in detail by Glaser.[19] As has already been seen in section 15·3, the focal length of this field is given by

$$\frac{a}{f_n} = (-1)^{n-1} \sin \frac{n\pi}{(1 + k^2)^{\frac{1}{2}}} \qquad [17\cdot31]$$

with

$$k^2 = \frac{eH_o^2a^2}{8mc^2\Phi} \qquad [17\cdot32]$$

n is the order of focus ($n - 1$ the number of intermediate images formed), so that under normal conditions of imaging $n = 1$. The formula obtained by Glaser for the spherical aberration of this lens is, for large magnification M,

$$\frac{Cf}{a} = \frac{n\pi}{4}\frac{k^2}{(k^2 + 1)^{\frac{3}{2}}}\frac{1}{\sin^4 \frac{n\pi}{(k^2 + 1)^{\frac{1}{2}}}} - \frac{1}{4}\frac{4k^2 - 3}{4k^2 + 3}\frac{\cos \frac{n\pi}{(k^2 + 1)^{\frac{1}{2}}}}{\sin^3 \frac{n\pi}{(k^2 + 1)^{\frac{1}{2}}}} \qquad [17\cdot33]$$

Figure 17·13 compares the spherical aberration of the bell-shaped field (Eq. 17·30) with that of System C (Fig. 17·3), the half-value widths of the two fields being adjusted so that their focal lengths become the same for low values of the field at the center of the lens. Since under these circumstances Eq. 17·31 yields (with $n = 1$)

$$\frac{a}{f} = \frac{\pi eH_o^2a^2}{16mc^2\Phi} = \frac{0.0345a^2H_o^2}{\Phi} \qquad [17\cdot34]$$

and the focal length of the System C is given by

$$\frac{D}{f} = \frac{0.01114H_o^2D^2}{\Phi} \qquad [17\cdot35]$$

[19] See reference 18.

the focal lengths of the two lenses will be the same for weak fields if the diameter D of the pole pieces of System C is made equal to $3.09a$. The half-value width $2a'$ of the field distribution of System C is then given by

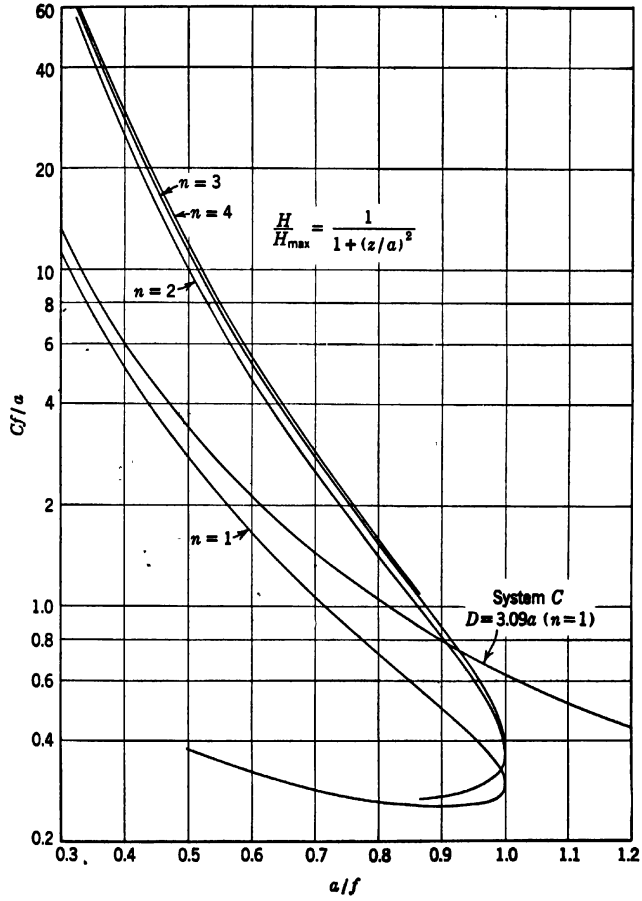


FIG. 17·13. Spherical Aberration as Function of Refractive Power for the "Bell-Shaped" Field (n = 1, 2, 3, 4) and for System C in Fig. 17·3 (n = 1).

$a' = 1.034a$. As might have been expected, the spherical aberrations of the two systems differ little, although greater refractive powers are attained with System C, whose field falls off more rapidly for large values of z.

As has been seen in section 15·3, a strong magnetic lens has a multiplicity of focal points. If the object is placed at the focal point farthest removed from the (highly magnified) final image, intermediate images will be formed at the intervening focal points. If this is applied specifically to the bell-shaped field (Eq. 17·30), the expression for the focal length (Eq. 17·31) indicates that two focal points and corresponding focal lengths exist for $k^2 > 3$, three for $k^2 > 8$, four for $k^2 > 15$, etc. If a focal length of given order n is considered, this decreases first rapidly with increasing field strength (increasing k), attains a minimum equal to

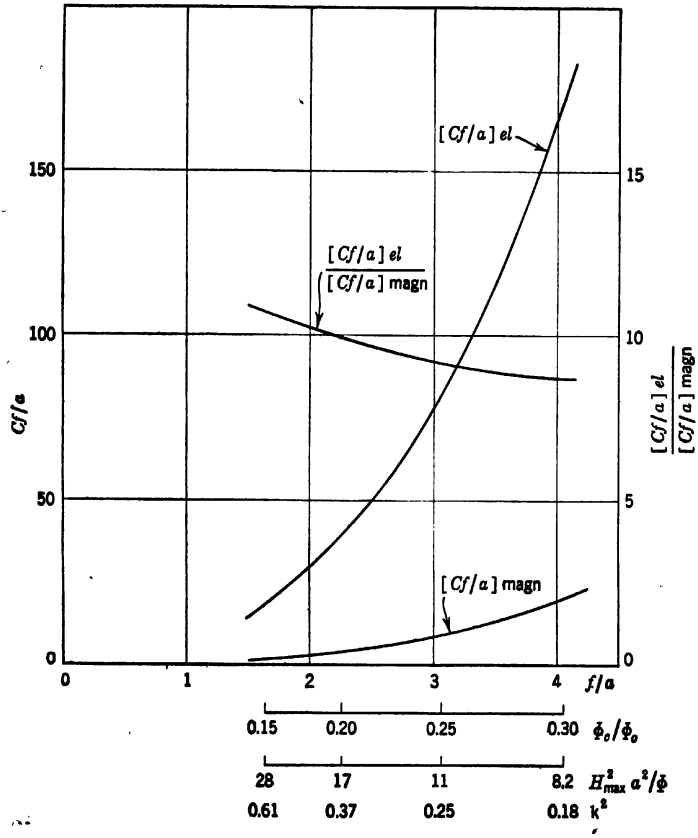


FIG. 17·14. Comparison of Spherical Aberrations of Magnetic and Electrostatic Unipotential Lens with Corresponding Field and Potential Distribution. (Dosse, reference 9.)

a at $k^2 = (2n)^2 - 1$, and then increases again relatively slowly. Figure 17·13 shows that, for any particular order n, the least spherical aberration is attained for a field stronger than that required to yield the minimum focal length. Throughout, the spherical aberration is smallest for the lowest order of imaging ($n = 1$ — no intermediate image), increasing with the number of intermediate images formed. However, for orders higher than the second the difference becomes negligible. A range of k^2 from 0 to 35 is covered by the plot.

Dosse[20] has also calculated the spherical aberration of an electrostatic unipotential lens with the axial potential distribution

$$\Phi = \Phi_o \left\{ 1 - \frac{1 - \frac{\Phi_c}{\Phi_o}}{1 + \left(\frac{z}{a}\right)^2} \right\} \qquad [17\cdot36]$$

and compared it with that of the analogous magnetic lens (Eq. 17·30). Φ_c is seen to be the axial potential at the center of the lens. Figure 17·14 shows the results of the calculation. For the short focal lengths considered, the aberration constants for the electric lens are approximately ten times as great as those of the magnetic lens.

In another investigation Dosse[21] determined the effect on the lens properties of rendering the refractive field asymmetric. By varying the ratio a_1/a in the field

$$H = \frac{H_o}{1 + \left(\frac{z}{a_1}\right)^2} \quad z < 0 \qquad H = \frac{H_o}{1 + \left(\frac{z}{a}\right)^2} \quad z > 0 \qquad [17\cdot37]$$

Dosse found that the spherical aberration decreased as a_1 was made larger, increased as it was made smaller than a. Thus the most favorable condition (for fixed H_o and a) consisted in a uniform field $H_o(a_1/a = \infty)$ continuing the bell-shaped field in object space, the least favorable condition ($a_1/a = 0$), in an abrupt drop of the field from H_o to zero. In the latter case the spherical aberration becomes infinite. For $k^2 > 3$ the field to the left of the origin becomes, of course, indifferent (Fig. 17·15).

In addition to these fields, the spherical aberration can generally be obtained analytically for any field for which r (in the case of a magnetic field) or $R = r\Phi^{1/4}$ (in the case of an electric field) is prescribed in the form of a simple function. For, in this case, the magnetic and electric fields

[20] See reference 9.

[21] See reference 19.

are similarly given by a more or less simple function. For a magnetic field

$$H = \left(-\frac{8mc^2\Phi}{e}\frac{r'}{r}\right)^{1/2} \qquad [17\cdot38]$$

and for an electric field

$$\Phi = \Phi_0 \exp\left\{\pm\int_0^z -\left(\frac{16}{3}\frac{R'}{R}\right)^{1/2} dz\right\} \qquad [17\cdot39]$$

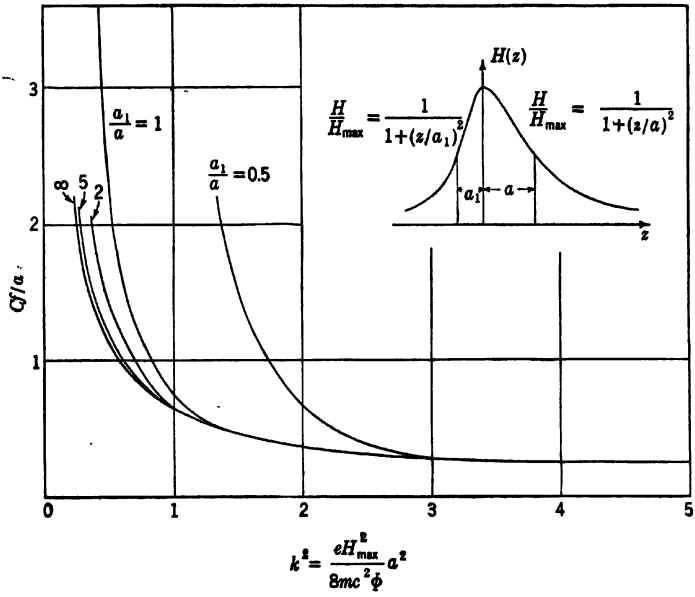


FIG. 17·15. Spherical Aberration of Bell-Shaped Field for Various Degrees of Asymmetry. (Dosse, reference 19.)

This circumstance has been utilized by Rebsch[22] to establish a practical lower limit for the spherical aberration of an electron-microscope objective. Rebsch argues that the rate of change of the refractive fields is limited by the precision of workmanship of which a mechanic is capable. Thus a definite lower limit is fixed for the distance l, where $(H'/H)_{max} = 1/l$ in the case of a magnetic and $(\Phi''/\Phi')_{max} = 1/l$ in the case of an electric field. Furthermore, he deduces from Eq. 17·1 that, for fixed l, those fields will give the least spherical aberration, for

[22] See reference 2.

which the fields themselves and the rates of change of the fields are a maximum right near the object (where the factor r_a^4 is smallest), falling off rapidly at larger distances. Since the spherical aberration, for fixed aperture angle θ and fixed (large) magnification, is proportional, according to Eq. 17·2, to Cf, it is proportional also, for given field distribution, to the scale of the objective, that is, also to l, the coefficient characteristic of the system being accordingly Cf/l. The question, then, is how small can Cf/l be made? Rebsch endeavored to answer this question by examining quite a number of fields having the property of causing rays leaving the point $z = 0$ on the axis to leave the field parallel to the axis in addition to having their maximum or maximum rate of change at this point. His results are tabulated in Table I. The fields are described in each case by the function $T(z)$, which is given by $eH^2/(8mc^2\Phi)$ for magnetic and by $(3/16)(\Phi'/\Phi)^2$ for electric lenses. J_0 and J_1 represent the Bessel function of zero and first order, respectively, $a = 2.40$ being the first root of the former.

Two deductions can be made from the figures listed in the table. First, for fields of this general type the quantity l and the focal length f are practically equal, so that $Cf/l \cong C$. Second, a probably general lower limit for C for practical systems forming real images is given by

$$C \geq 0.25 \qquad [17\cdot40]$$

The existence of smaller values of C has, it is true, been demonstrated by

TABLE I

System	$T(z)$	$R(z)$	Cf/l	l/f
r. magn.	a^2e^{-2z}; $a = 2.40$	$J_0(ae^{-z})/[aJ_1(a)]$	0.252	1.25
s. electr.	a^2e^{-2z}; $a = 2.40$	$J_0(ae^{-z})/[aJ_1(a)]$	0.388	1.10
t. magn.	$3/(1+z^2)^2$	$z/(1+z^2)^{1/2}$	0.294	1.00
u. electr.	$3/(1+z^2)^2$	$z/(1+z^2)^{1/2}$	0.438	1.20
v. magn.	$\dfrac{4ze^{-2z}}{1-(1+z)e^{-2z}}$	$1-(1+z)e^{-2z}$	0.311	1.00
w. magn.	$4e^{-z}/(3-e^{-z})$	$3/2 - 2(e^{-z} - e^{-2z}/4)$	0.324	0.889
x. magn.	$2\,\mathrm{sech}^2 z$	$\tanh z$	0.410	1.00

Glaser,[23] who showed that, if in System *t* the field strength is increased by a factor 3.4 and the object pushed correspondingly beyond the center of the field, C is reduced to about 0.18. However, this is not a particularly favorable point of operation, since $Cf/l = 0.37$. The latter quantity nowhere is less than 0.252.

[23] See reference 18.

It is interesting to note that the exponential field, System r, which was shown by Glaser[24] to form the optimum short lens, leads also here to the least aberration of all the systems examined.

It has been pointed out on previous occasions that the aperture defect can be nullified completely if electron mirrors are employed in conjunction with electron lenses. This can be demonstrated very simply by considering an electron lens terminated in a retarding field (Fig. 17·16). Assume that the lens, considered as terminated at the plane AA, has a

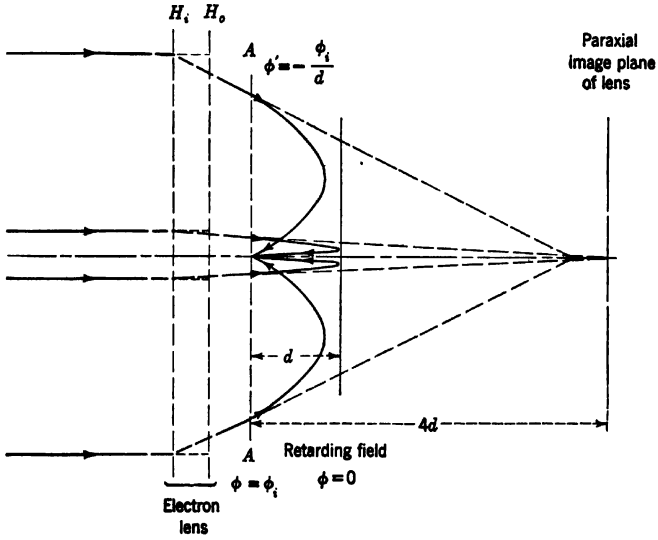


FIG. 17·16. Correction of Spherical Aberration of an Electron Lens by Combination with a Retarding Field Acting as an Electron Mirror.

virtual paraxial focus at the distance $4d$ from AA, as indicated by the dotted lines representing the tangents to the electron rays at AA, d being the distance within which the potential is reduced from Φ_i to 0. Furthermore, assume that the spherical aberration of the lens is given by $\Delta r_i = -4d\theta^3$. This is a perfectly reasonable value for the spherical aberration of the lens, corresponding to $C \cong 1$. Then, by Eq. 13·54, the rays are focused, exactly to the third order of approximation, at the intersection of the axis with the plane AA. The system described thus represents an electron-optical system fully corrected for spherical aberration.

[24] See reference 10.

The above system has the drawback that the image or object, whichever is placed at A, is located in a strong electric field. If the object is placed at A (as would be the case if the system were to be used as an electron-microscope objective) it is furthermore necessary to illuminate the object from the lens side with a narrow pencil of electrons. It is difficult to do so without cutting off either a large part of the useful lens aperture or, if the electron source is placed farther away from the lens, without cutting off a part of the image and using a decentered beam for illumination. Similar difficulties arise with most arrangements utilizing an electron mirror or mirror-lens combination as a short-focus electron objective.

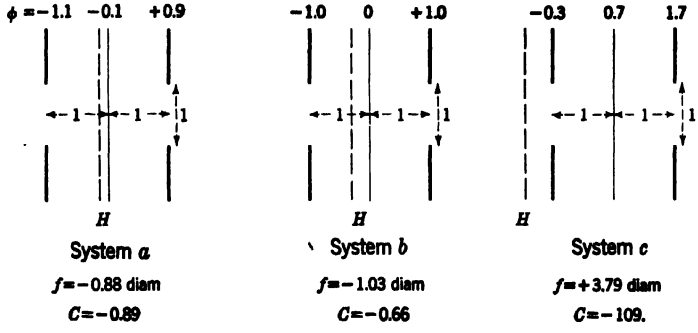


FIG. 17·17. Three Simple Electron Mirrors and Their Aperture Defects.

Very little has been done in the determination of the magnitudes of the aberrations of electron mirrors. This is not surprising in view of the relative complexity of the formulas. Figure 17·17 shows three simple electron mirrors, indicating their cardinal points and the spherical aberration constant C. Systems a and b, both representing "convex" mirrors, have values of C (-0.89 and -0.66, respectively), which are small and of the wrong sign to make them useful for correcting the aperture defect of ordinary electron lenses. It should be remarked that in combining such mirrors with electron lenses, it is generally necessary to combine them with two similar lenses, one ahead of and one behind the mirror. The geometrical situation is such that the electrons necessarily traverse the lens twice. With "concave" mirrors, such as System c, which has a very large spherical aberration ($C = -109$) opposite in sign to that of an electron lens, it should be possible to obtain a corrected image placing the object between the mirror and the lens and closer to the mirror than its focal point, so that the refraction is divided between lens and mirror (Fig. 17·18).

Another basic possibility of correcting electron lenses for spherical aberration is the employment of ultra-high-frequency fields to modify the refractive power of an electron lens in the interim between the arrival of the paraxial and the marginal electrons at the lens. Since these intervals are very small, the frequencies required are very high. Further-

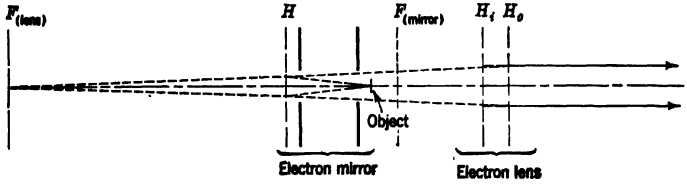


FIG. 17·18. Combining a Converging Electron Mirror and an Electron Lens to Form a System Free from Aperture Defect. (Object Illuminated from the Right, through the Lens.)

more, if the electrons are to participate in the formation of the image for more than an infinitesimal fraction of the ultra-high-frequency cycle, it is necessary that their velocity be modulated in synchronism with the lens fields, maintaining the refractive power of the lens constant during the useful part of the cycle.

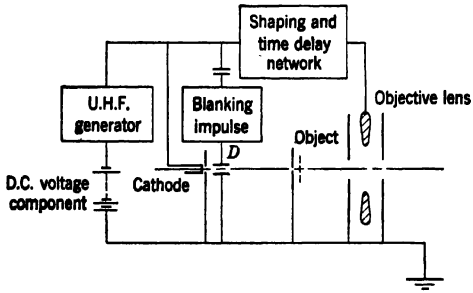


FIG. 17·19. Schematic Circuit for Correcting Spherical Aberration by Simultaneous Modulation of Electron Velocities and Lens Voltage.

Figure 17·19 shows a schematic circuit diagram of an electrostatic objective spherically corrected in this manner. The anode and outer electrodes of the electron lens are grounded. A properly shaped voltage wave is applied to the cathode and another, in general differently shaped, wave to the center element of the electron lens. Between the anode and the object the path is kept short and the beam collimated with great care. A deflection system D may be provided to " blank out " the beam

during a portion of the cycle. Figure 17·20 shows the voltage waves that would have to be applied to the cathode and to the center element of the objective for the special case that the distance between the anode and the lens $l = 1$ centimeter, that the focal length $f = 0.5$ centimeter, the spherical-aberration coefficient of the lens is $C = 10.5$, and the chromatic-aberration coefficient $C_3 = -4.1M\Delta\Phi/\Phi$, the lens being treated, not quite properly, as a thin lens. The dotted portions of the waves are

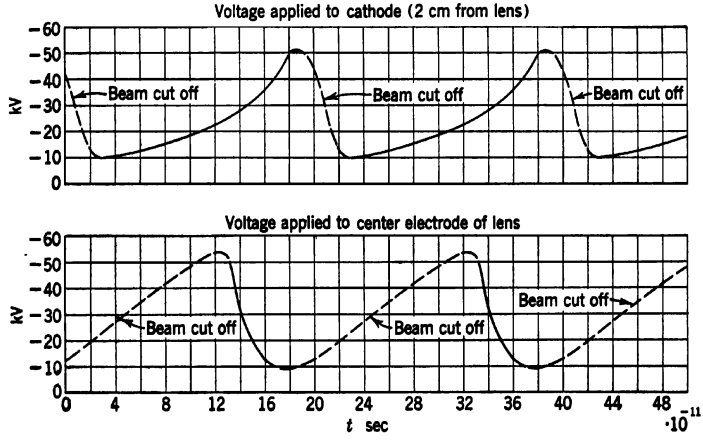


FIG. 17·20. Voltage Waves to Be Applied to Cathode and Center Electrode of Objective in System Corrected for Spherical Aberration by High-Frequency Modulation (Fig. 17·19).

arbitrary and correspond to the fraction of the cycle during which the electron beam is cut off by the synchronized deflection system. It is seen that the frequency is extremely high, the fundamental wave length being 6.0 centimeters, so that the precise shaping required, among other things, presents considerable difficulties. It should also be mentioned that this dynamic method of correcting spherical aberration, although very attractive in many ways, demands unusual uniformity in the initial velocity of the electrons.

17·3. Coma. The magnitude of the coma defect is determined by the expressions for S_6 and S_7, as given by Eqs. 16·64*f* and 16·64*g*. Both coefficients, determining the contribution of isotropic and anisotropic coma, respectively, are seen to vanish if both the electric field Φ' and the magnetic field H are antisymmetric about some point $z = z_m$, at which the aperture is placed, and if the magnification is -1. For in this case Φ, Φ'', H', r_α, r'_γ are symmetric and Φ', H, r'_α, r_γ antisymmetric

about $z = z_m$, so that every term under the integral sign becomes antisymmetric, making the integral itself zero.[25]

For a weak, short lens (with the aperture plane in the middle of the lens) Eq. 16·64*f* becomes:

$$S_6 = \frac{M}{16 r_{aa}^2 \Phi^2} \int_{z_o}^{z_i} \left\{ \frac{5\Phi''^2}{4} + \frac{eH'^2\Phi}{mc^2} \right\} r_\gamma r_\alpha^3 \, dz \qquad [17\cdot41]$$

where r_α may be regarded as constant and equal to r_{aa}. If the lens fields are symmetrical as before, r_γ will be antisymmetric about the plane of symmetry, no matter what the magnification. Hence any weak short lens with antisymmetric electric and magnetic fields with the aperture placed at the center of the lens is free of isotropic coma.[26]

For a purely magnetic lens, the anisotropic coma coefficient S_7 takes on the following simple form:

$$S_7 = M \left(\frac{eH_o^2}{8mc^2\Phi} \right)^{1/4} \int_{z_o}^{z_i} \left\{ \frac{eH_o^2}{8mc^2\Phi} \left(\frac{H}{H_o} \right)^3 + \frac{H''}{8H_o} \right\} \frac{r_\alpha^2}{r_{aa}^2} \, dz \qquad [17\cdot42]$$

Thus an antisymmetric short magnetic lens quite generally is free of anisotropic coma.

Voit,[27] furthermore, has shown that an arbitrary lens field with arbitrary position of the aperture can be freed from isotropic coma by a slight modification of its electric or (scalar) magnetic potential without altering its first-order properties. The procedure consists in introducing a " kink " in the curve for the axial potential, that is, a near-singularity of the second derivative of the potential (Φ'' or H'), at a convenient point without greatly altering either the potential or its first derivative in the neighborhood. It is true that this prescription has more theoretical than practical significance.

In general, coma is of small importance in electron optics. In electron guns, concerned only with the formation of a narrow central spot without regard to faithfulness of reproduction, it is obviously of no moment. Again, in devices such as the image tube, which serve to image a large area with narrow pencils, it is of no significance in comparison with curvature of field and astigmatism, which increase more rapidly with the separation of the object point from the optic axis. It might be expected to play a larger role in connection with the electron microscope, since it is known to be one of the aberrations which require careful correction in light-microscope objectives.

A calculation for a characteristic strong magnetic objective, such as

[25] See Scherzer, reference 1.
[26] See Riedl, reference 20.
[27] See reference 21.

the system consisting of two cylindrical pole pieces with a narrow gap (System C), with $H_{max}^2/V = 150(2.630/D)^2 = 1040/D^2$ gauss^2/volt ($f = 0.261D$), where D is the clear diameter of the pole pieces in centimeters, suggests that here also it is of little importance — at least as long as spherically uncorrected electron lenses are employed, as at present. Thus, if the effective aperture is placed in the image-side focal plane, that is, is considered as being determined by the size of the (distant) source of illumination as viewed from the object, it is found that S_6/M takes on the value $-14.5/D^2$, whereas S_7/M is equal to $9.36/D^2$. The maximum diameter of the aberration figure (see Eqs. 16·30 and 16·31) is $3(S_6^2 + S_7^2)^{1/2} r_o r_a^2 = 52 M r_o r_a^2/D^2$, r_o being the separation of the object point from the axis and r_a the radius of the effective aperture. Assume that this maximum dimension of the aberration figure is to be nowhere greater than $50 \cdot M$ A.U. and consider the special case that the effective aperture angle is $\theta = 3 \cdot 10^{-3}$ radian, and the magnification is $M = 10{,}000$. The permissible image diameter becomes then

$$2Mr_o = \frac{10^{-2}}{52 f^2 \theta^2} = 313 \text{ cm} = 10\tfrac{1}{2} \text{ ft}$$

Obviously, quite generally, coma is not one of the important image defects of the electron microscope.

17·4. Curvature of Field and Astigmatism. Curvature of field and astigmatism are closely related insofar as they both depend on the first power of the aperture radius and the second power of the separation of the object point from the axis. From an analysis of the expressions for S_3, S_4, and S_5 (Eqs. 16·64c–16·64e) Voit[28] derives the following conclusions:

1. For a purely magnetic lens S_3/M is always positive, since the terms under the integral can be arranged as a sum of squares. Consequently, for a flat object the real image is invariably concave toward the lens.

2. For an electric short lens S_3/M, similarly, is always positive, so that here also the image is concave toward the lens. However, it is possible to prescribe, for an electric lens of given magnification, certain field modifications near the object which will nullify the curvature of the lens. Since these field modifications can probably not be realized in practice, the curving of the object surface may, here also, be the only practical way of achieving a flat image.

3. Isotropic astigmatism (S_4) and anisotropic astigmatism (S_5) can simultaneously be made zero even for the weak magnetic lens by a

[28] See reference 21.

choice of the aperture plane and a sufficient number of disposable constants in the expression for the magnetic field.

4. Astigmatism (S_4) can be made zero for a weak as well as for a strong electric lens by suitable choice of the aperture plane and the field distribution.

Curvature of field and astigmatism are significant only for lenses imaging extended areas. Thus they are of no importance in electron guns (unless deflection takes place ahead of or at the lens). In the electron microscope, objective and projector contribute equally to curvature and astigmatism in the image. Let the total magnification be $M = M_1 \cdot M_2$, where the subscript 1 refers to the objective and the subscript 2 to the projector. Then, for example, curvature of field gives rise to an aberration disk of radius Δr_i, y_o' being the coordinate of the object point:

$$\Delta r_i = M_2 S_{31} y_{o1}'^2 r_{a1} + S_{32} y_{o2}'^2 r_{a2} = M_2 S_{31} y_{o1}'^2 f_1 \theta_o + S_{32} y_{o1}'^2 M_1^2 f_2 \frac{\theta_o}{M_1} \qquad [17\cdot43]$$

$$= M \left(\frac{S_{31}}{M_1} f_1 + \frac{S_{32}}{M_2} f_2 \right) y_o'^2 \theta_o$$

Thus, if the two lenses are similar in character and have the same focal lengths, their effect on the curvature and astigmatism of the image is identical. As in the case of coma, the aberration coefficients in question have been calculated for the two-cylinder magnetic lens with $H_{max}^2/V = 1040/D^2$ gauss^2/volt ($f_1 = 0.261D$). The values obtained are

$$\frac{S_3}{M} = \frac{80.6}{D^2} \qquad \frac{S_4}{M} = \frac{21.0}{D^2} \qquad \frac{S_5}{M} = -\frac{19.4}{D^2}$$

In addition, they were calculated for a similar magnetic lens with $H_{max}^2/V = 470/D^2$ gauss^2/volt ($f_2 = 0.385D$) used as a projector. At this point the projector magnification of the lens reaches a maximum (Fig. 4·19). For this lens

$$\frac{S_3}{M} = \frac{8.75}{D^2} \qquad \frac{S_4}{M} = -\frac{3.78}{D^2} \qquad \frac{S_5}{M} = \frac{3.62}{D^2}$$

The maximum diameter of the aberration figure is given according to Eqs. 16·26, 16·28, and 16·29 by

$$2M \left\{ \left(\frac{S_{31}}{M_1} + \frac{S_{32} f_2}{M_2 f_1} \right) + \left[\left(\frac{S_{41}}{M_1} + \frac{S_{42} f_2}{M_2 f_1} \right)^2 + \left(\frac{S_{51}}{M_1} + \frac{S_{52} f_2}{M_2 f_1} \right)^2 \right]^{1/2} \right\} r_o^2 r_a$$

$$= 227 M r_o^2 r_a$$

if all lengths are measured in lens diameters. For an electron microscope having the stronger lens as objective, the weaker (geometrically identical) lens as projector, the diameter of the image over which the diameter of the aberration figure is less than 50 A.U. can be determined as before: With a total magnification of 10,000, an objective aperture of $3 \cdot 10^{-3}$ radian, and an objective focal length of 0.5 centimeter,

$$2Mr_o = 2M \cdot 1.92 \left[\frac{5 \cdot 10^{-7}}{227 \cdot 0.5 \cdot 3 \cdot 10^{-3}} \right]^{1/2} = 46.5 \text{ cm}$$

Thus with this rather stringent requirement the diameter of the image may still be over 1½ feet. If the object is made concave toward the objective, being given a radius of curvature of the order of 0.5 millimeter, the curvature of the image is nullified, permitting twice as large an image diameter for the same requirement with regard to resolution. It is seen that in the electron microscope, also, curvature and astigmatism are of secondary importance.

This is not the case, however, for image tubes, that is, devices which focus an extended electron image from a cathode (for example, a photocathode on which a light image has been projected) onto some other surface, such as a fluorescent screen, at a higher potential. As has been pointed out before in connection with the general consideration on the aberrations of cathode lenses, the apertures of the imaging beams are in this case very small — depending on the initial lateral velocity components of the electrons — and spherical aberration and coma are consequently insignificant.

Calculations and measurements of the curvature of field and astigmatism have been made for a particular electrostatic image tube of approximately unity magnification, consisting of two equidiameter cylinders, the cathode cylinder being 0.9 and the anode cylinder 1.8 cylinder diameters in length.[29] With the cathode in the form of a flat disk, the tangential and sagittal image surfaces take on the form shown to scale in Fig. 17·21. The radii of curvature of these two surfaces at the optic axis were 0.022 and 0.038 diameter, respectively, corresponding to the values of the aberration coefficients S_3 and S_4:

$$S_3 = \frac{M^2}{4Z}\left(\frac{1}{R_t} + \frac{1}{R_s}\right) = -\frac{11.5}{D^2} = -\frac{17.7}{r_a D}\left(\frac{\Phi_{ro}}{\Phi_i}\right)^{1/2}$$

$$S_4 = \frac{M^2}{4Z}\left(\frac{1}{R_t} - \frac{1}{R_s}\right) = -\frac{3.08}{D^2} = -\frac{4.7}{r_a D}\left(\frac{\Phi_{ro}}{\Phi_i}\right)^{1/2}$$

Here $Z = 1.54D$ is the distance between the aperture plane, located at

[29] See Morton and Ramberg, reference 22.

the apparent point of divergence of the principal rays from the axis in image space, and the image plane. $e\Phi_{ro}$ is the initial kinetic energy in a lateral direction of the electrons.

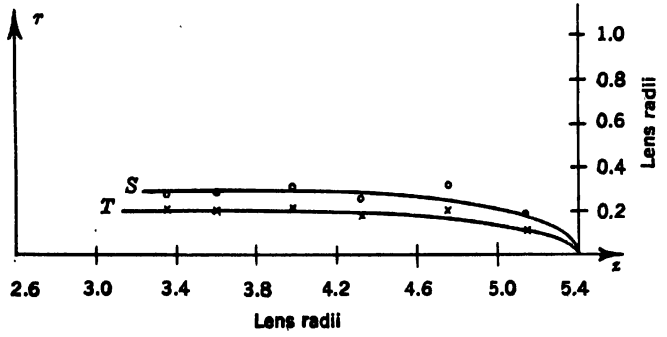


FIG. 17·21. Curvature of Image Field: *S*, Calculated Sagittal Image Surface; *T*, Calculated Tangential Image Surface; *o*, Measured Sagittal Image Points; *x*, Measured Tangential Image Points. (Courtesy of *Physics*, reference 22.)

The curvature of the image surfaces is great enough to render the image obtained definitely unsatisfactory, in spite of the small aperture of the imaging pencils. Figure 17·22 shows, on the left, the appearance

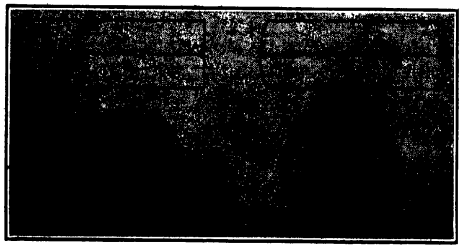


FIG. 17·22. Comparison of Images from Tubes with Flat Cathode (Left Side) and with Curved Cathode (Right Side). (Courtesy of *Physics*, reference 22.)

of a square mesh pattern imaged by a system of this type. On the right is reproduced the same pattern, when imaged by a similar system, whose cathode has been made concave so that its center of curvature comes to lie in the plane separating the two cylinders. The curving of the cathode in this case not only changes the object distances of the marginal object points, but alters both the refracting fields and the effective

position of the aperture plane materially (Fig. 17·23). The change is in the direction of causing the field conditions surrounding each principal ray to approximate closely those surrounding the optic axis. Thus imaging conditions off and on the axis come to resemble each other, and the field aberrations, curvature, astigmatism, and distortion, are made small.

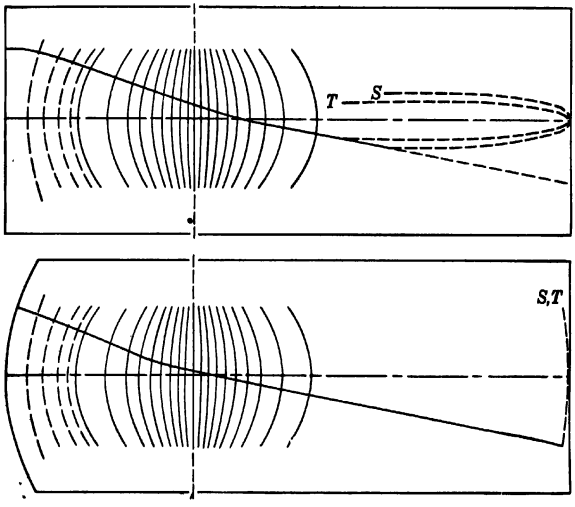


FIG. 17·23. Potential Distribution and Image Curvature in Tubes with Flat Cathode and with Curved Cathode (Schematic).

17·5. Distortion. For distortion the same general statements hold as for coma. Antisymmetric lens fields with the aperture plane at the center of symmetry are free of distortion for unity magnification of the image. If the antisymmetric fields are weak in addition, they are free of distortion for arbitrary magnification. Furthermore, for arbitrary field distributions the isotropic distortion can be removed by inserting a " kink " in the potential distribution at a suitable point.

In the case of the electron microscope only the distortion produced by the projector need, normally, be considered, since the isotropic distortion produced by both lenses is given by

$$\Delta r_i = M_2 S_{11} r_o^3 + S_{12} M_1^3 r_o^3 = M \left(\frac{S_{11}}{M_1} + M_1^2 \frac{S_{12}}{M_2} \right) r_o^3$$

the (last) subscript 1 referring once more to the objective, the subscript 2, to the projector.

For the magnetic two-cylinder lens with $H^2_{max}/V = 470/D^2$ gauss^2/volt ($f_2 = 0.385D$), the value of the coefficient of isotropic distortion, S_1/M, is $2.19/D^2$, the distortion being, accordingly, pincushion-shaped. The effective aperture plane lies, of course, at the image-side focal point of the lens. Suppose that the maximum tolerable distortion corresponds to a displacement Δr_i from the Gaussian image point equal to one-tenth of the distance r_i, the distance of the image point in question from the axis. Then

$$0.1Mr_o \geq M \cdot M_1^2 \frac{S_1}{M_2} \cdot r_o^3 \quad \text{or} \quad r_i^2 = M^2 r_o^2 \leq \frac{M_2^2}{10\dfrac{S_1}{M_2}}$$

so that for $M_2 = 100$, $r_i \leq 21.3D$. Thus, if the focal length of the projector is 0.50 centimeter, the diameter of the image may be at most $2 \cdot 1.30 \cdot 21.3 = 55.4$ centimeters.

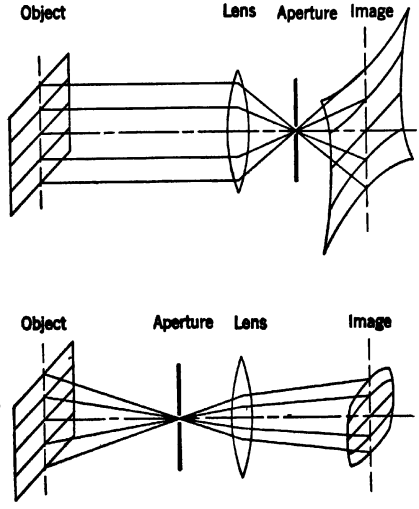


FIG. 17·24. Effect of Aperture Position on Character of Distortion.

In addition to isotropic distortion, the lens just considered has anisotropic or rotational distortion, the coefficient S_2/M being $6.5/D^2$ — actually much larger than the isotropic distortion. The effect of this rotational distortion is that the displacement from the Gaussian image point is over three times as great as for the isotropic distortion alone and inclined by an angle arctan $(S_2/S_1) = 71$ degrees to the radial direction. It restricts the useful image diameter to approximately 32 centimeters.

If, as in the example considered, the objective is employed with the object very close to the geometrical center of the lens, the distortion introduced by the objective may further limit the useful field.

The manner in which distortion is corrected by curving the cathode of an electrostatic image tube has been demonstrated in the last section (Figs. 17·22 and 17·23). This may be regarded as a special instance of a general rule, applying strictly for a thin lens. If the aperture plane (plane in which the principal rays of the imaging pencils intersect the axis) is ahead of the lens, barrel-shaped distortion results; if it is beyond the lens, cushion-shaped distortion results. This is a consequence of the normal spherical aberration of electron lenses, as is shown in Fig. 17·24. In the great majority of cases the effective aperture is beyond the lens and the distortion is, correspondingly, cushion-shaped. If it is possible, by some means, such as the curving of the cathode in an image tube, to approach the aperture plane closer to the effective center of the lens, a reduction in the distortion may be attained.

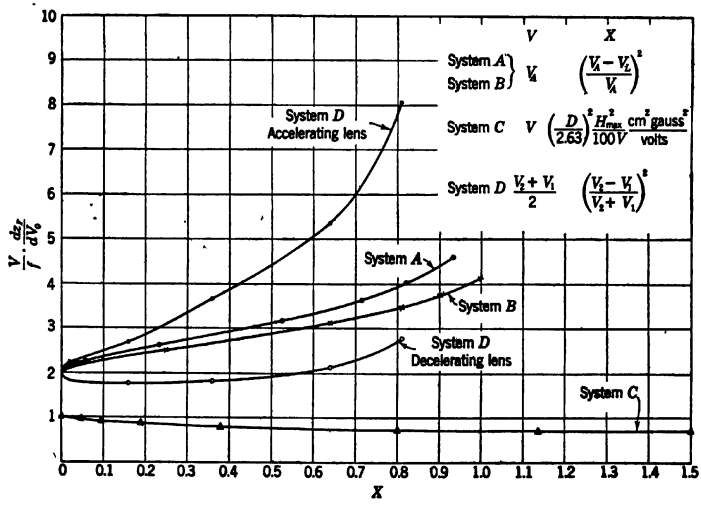


FIG. 17·25. Variation of Chromatic Aberration of Systems *A–D* (Fig. 17·3) with Lens Strength. (Courtesy of *J. Applied Phys.*, reference 8.)

17·6. Chromatic Aberration. The impossibility of correcting the axial chromatic aberration of an electron lens, given by the coefficient C_3, Eqs. 16·102 and 16·103*c*, was demonstrated in section 16·6. Similarly, certain upper limits for the magnitude of this coefficient were established for purely magnetic and purely electric lenses. Figure 17·25 goes

beyond this point by showing the variation of the coefficient $-(C_3/M)(\Phi/\Delta\Phi) = (dz_F/d\Phi)(\Phi/f)$, where Φ denotes the object- and image-space potential in the case of the unipotential lenses and the mean potential in the case of the immersion lens, for the four typical electron lenses Systems A, B, C, and D (Fig. 17·3) discussed in section 17·2, with the parameters determining the strength of the lenses as abscissas. The object is in all cases assumed to lie close to the focal point of the lens.

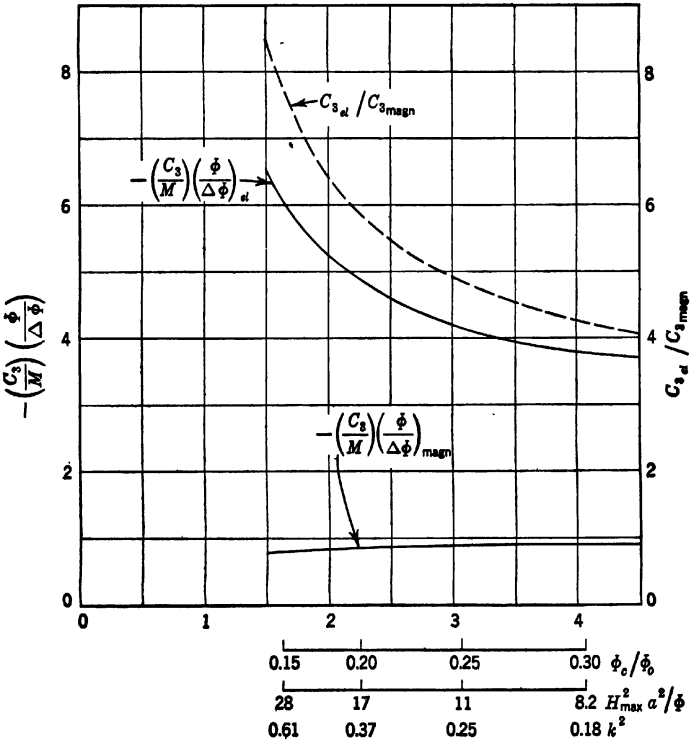


FIG. 17·26. Comparison of Chromatic Aberrations of Magnetic Lens and Electrostatic Unipotential Lens with Corresponding Field and Potential Distributions. (Dosse, reference 9.)

It is seen that the quantity $-(C_3/M)(\Phi/\Delta\Phi)$ increases uniformly from the value 2 for the weak lens in the case of the electrostatic lenses with negative center electrode (Systems A and B) and decreases uniformly from the value unity for low lens strengths in the case of the

magnetic lens (System C). This is in accord with the upper limits established in section 16·6. For the immersion lens used as a decelerating lens the same quantity decreases first, then increases at higher values of the lens strength. When it is used as an accelerating lens, this quantity increases uniformly with the lens strength. It is evident from these curves that the chromatic aberration of a magnetic lens is in all cases considerably lower than that of an electrostatic unipotential lens with negative center electrode. This is also brought out by the curves in Fig. 17·26, which represent Dosse's calculations of the chromatic aberration of a magnetic and an electrostatic unipotential lens with the field distributions $H = H_o/(1 + [z/a]^2)$ and $\Phi = \Phi_o - (\Phi_o - \Phi_c)/(1 + [z/a]^2)$, respectively.[30] If, in an electrostatic unipotential lens, the center electrode is made positive with respect to the outer ones, the upper limit prescribed by Eq. 16·119 requires that $-(C_3/M)(\Phi/\Delta\Phi)$ be less than 2. Accordingly, this quantity may be calculated to be equal to 1.42 for System A with the center electrode at approximately twice the potential of the other electrodes. Its refractive power is in this case $1/f = 0.0889/D$.

It has already been pointed out that the chromatic aberration of an electron lens can be compensated by combining the lens with an electron mirror. From the fact that the converging action of an electrostatic mirror increases quite generally when its negative electrode is made less negative it follows that the faster electrons, penetrating deeper into the mirror field, are refracted more strongly by the field than the slower electrons. The sign of the chromatic aberration of the electron mirror is normally opposite to that of the electron lens. Figure 17·27 shows the calculated variation of the coefficient with the ratio of the potentials of the electrodes for an electrostatic mirror consisting of two equidiameter cylinders, the potential of the negative electrode being throughout sufficiently low so that the mirror acts as a converging system. Its refractive power, $1/f$, is also indicated on the graph. It is seen that, as in the case of spherical aberration, the aberration coefficients of the converging or " concave " electron mirror are considerably larger than for electron lenses of corresponding dimensions and focal lengths.

A particularly close relation between spherical and chromatic aberration exists in the case of the superposed uniform electric and magnetic field. Here the kinetic energy of the lateral component of velocity of any electron is a constant, and its magnitude is without effect on the axial point P' at which electrons leaving the point P on the axis with the same axial velocity converge again. Spherical aberration arises only from the fact that for electrons with the same total velocity the axial

[30] See Dosse, reference 9.

component must decrease as the lateral component is increased. As chromatic aberration arises from the same source, that is, a variation in the axial velocity for a fixed lateral velocity, either aberration can

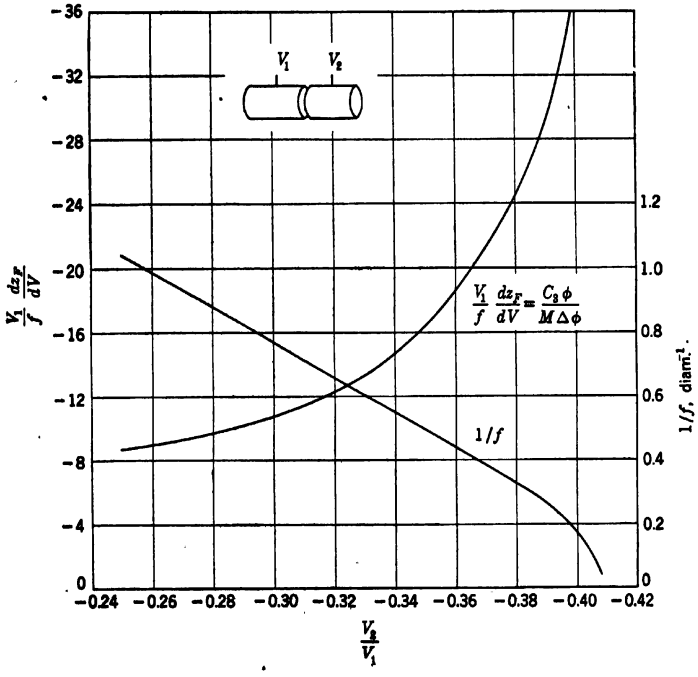


FIG. 17·27. Axial Chromatic Aberration of Converging Electron Mirror.

readily be deduced from the other. Thus, from Eq. 17·27, since $\theta_o^2 = \Phi_r/\Phi_o = -\Delta\Phi/\Phi_o$, the chromatic aberration of a uniform magnetic field is given by

$$\Delta r = \frac{10.54}{H}\frac{\Delta\Phi}{\Phi^{\frac{1}{2}}}\theta_o \qquad [17·44]$$

and that of a superposed electric and magnetic field, by

$$\Delta r = \frac{10.54}{H}\frac{\Delta\Phi}{\Phi_i^{\frac{1}{2}}}\theta_o \qquad [17·45]$$

For a uniform magnetic field superposed on a uniform electric mirror field the sign of the aberration is reversed, the time of travel from the object to the image increasing instead of decreasing with the initial

velocity. If $L = -\Phi_o/\Phi'$ is the distance between the object and the plane of reversal of electrons with initial kinetic energy $e\Phi_o$, the chromatic aberration is given by

$$\Delta r = \frac{\Delta\Phi}{\Phi_o} L \left(1 + \left[\frac{\Phi_o}{\Phi_i}\right]^{1/2}\right) \theta_o = -\frac{10.54}{H} \frac{\Delta\Phi}{\Phi_i^{1/2}} \theta_o \qquad [17·46]$$

The chromatic aberration of cathode lenses, discussed in detail in section 16·8, cannot be corrected with the aid of an electron mirror, since the shift in the image plane for such lenses is, to the lowest order of approximation, proportional to $(\Delta\Phi/\Phi)^{1/2}$ whereas that of the mirror is proportional to $\Delta\Phi/\Phi$. It is possible, however, by combining two such

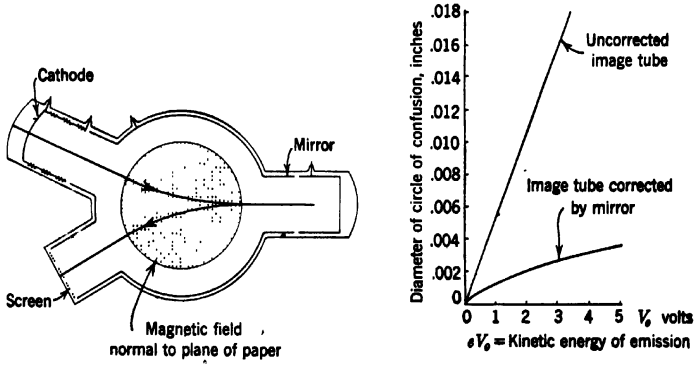


FIG. 17·28. Correction of the Chromatic Aberration of an Image Tube with an Electron Mirror.

electron-optical elements, to bring the electrons with two different initial axial velocities to focus at the same point so that, for a given spread in the initial velocities at the cathode, a considerable improvement in the image can be attained by the employment of a mirror. Figure 17·28 shows an image tube employing an electron mirror to obtain such correction together with a graph indicating the effect of the mirror on the focusing for different initial velocities. A uniform magnetic field normal to the plane of the paper prevents the reflected electrons from returning to the cathode.

The chromatic difference in magnification of electron lenses, given by the coefficient C_1 (Eqs. 16·102 and 16·103a), depends materially on the location of the effective aperture. For antisymmetric fields with the aperture in the plane of symmetry and with unity magnification this aberration vanishes. If, in addition to being antisymmetric, the lens is

weak, chromatic difference in magnification is zero irrespective of the magnification.

The corresponding anisotropic defect, chromatic difference in rotation, given by C_2 (Eqs. 16·102 and 16·103*b*), exists only for magnetic lenses. For purely magnetic lenses it is proportional to the image rotation produced by the lens:

$$C_2 = -\frac{M}{2}\frac{\Delta\Phi}{\Phi}\chi \qquad [17\cdot47]$$

where χ is the angle of image rotation. The two defects together cause a deviation of the image point from the Gaussian image point equal to $(C_1^2 + C_2^2)^{1/2} r_o$, r_o being the distance of the object point from the axis.

For the equidiameter cylinder magnetic lens (System *C*) with $H_{max}^2/V = 1040/D^2$ gauss^2/volt, C_1 and C_2 are given by

$$C_1 = 1.1M\frac{\Delta\Phi}{\Phi} \qquad C_2 = -1.04M\frac{\Delta\Phi}{\Phi}$$

so that the deviation from the Gaussian image point makes an angle of approximately 45 degrees to the radial direction. Since $C_3 = -0.695M(\Delta\Phi/\Phi)$, the condition that in the image the effect of the chromatic difference in magnification and rotation should not exceed that of the axial chromatic aberration:

$$(C_1^2 + C_2^2)^{1/2} r_o \leq |C_3| r_a$$

leads for $\theta = 3\cdot10^{-3}$, $f = 0.5$ centimeter, and $M = 10{,}000$ to a value for the radius of the image $r_i = Mr_o = 6.85$ centimeters.

For System *C* with $H_{max}^2/V = 470/D^2$ gauss^2/volt, corresponding approximately to the maximum useful coil current in a projection lens of this type,

$$C_1 = -0.19M\frac{\Delta\Phi}{\Phi} \qquad C_2 = -1.22M\frac{\Delta\Phi}{\Phi}$$

so that here the anisotropic defect predominates. For weaker lenses this is not the case, since C_1 approaches the value $-M(\Delta\Phi/\Phi)$, whereas C_2 decreases continuously in proportion to the field as the magnetic field is reduced to very small values.

If the stronger of the two lenses considered in detail is employed as objective, the weaker as projector, the chromatic defects add up to distort an image point at a distance r_i from the axis into a line of length

$$r_i\frac{\Delta\Phi}{\Phi}([-1.1 + 0.19]^2 + [1.04 + 1.22]^2)^{1/2} = 2.43r_i\frac{\Delta\Phi}{\Phi}$$

inclined at an angle arctan (2.26/0.91) = 68 degrees to the radial direction. By reversing the flow of current in one of the coils, the coefficient

is reduced from 2.43 to 0.93 and the angle with the radial direction to 11 degrees.

It has already been shown in section 16·8 that for simple types of cathode lenses the effect of chromatic difference in magnification and rotation is insignificant in comparison with axial chromatic aberration.

17·7. Space-Charge Defect. Methods of designing electrode structures so as to prevent the spreading of electron beams under the influence of space-charge forces have been worked out by Pierce.[31] As an example a space-charge limited, axially symmetric electron beam will be considered below.

Within the beam Poisson's equation

$$\frac{\partial^2\phi}{\partial z^2} = 4\pi j\left(\frac{m}{2e\phi}\right)^{\frac{1}{2}} \qquad [17\cdot48]$$

with the side condition

$$\frac{\partial\phi}{\partial r} = 0 \qquad [17\cdot49]$$

must hold; j is the current density within the beam. If the beam is space-charge limited, the gradient at the cathode must be placed equal to zero and the solution of Eq. 17·48 becomes

$$\phi = Az^{\frac{4}{3}} \qquad [17\cdot50]$$

with

$$A = (9\pi j)^{\frac{2}{3}}\left(\frac{2e}{m}\right)^{-\frac{1}{2}} = 5.69\cdot10^3 j^{\frac{2}{3}} \text{ volts per cm}^{\frac{4}{3}}$$

The conditions Eqs. 17·50 and 17·49 must hold on both sides at the boundary between the beam and the space-charge-free region surrounding it. This fixes the boundary conditions for the free region and, hence, the shape and potentials of the electrodes enclosing it. The actual determination of the electrode shapes is best made with the aid of an electrolytic wedge tank,[32] as shown in Fig. 17·29. The edge of the beam is represented in the tank by an insulating strip so as to satisfy the bound-

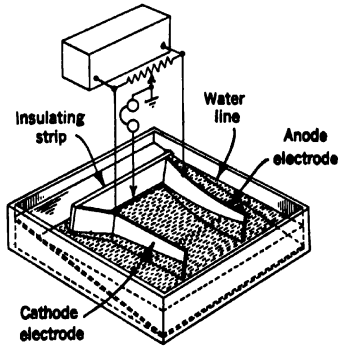


FIG. 17·29. Wedge Tank for Determining Electrode Shapes for Forming Parallel Space-Charge-Limited Beam. (Courtesy of *J. Applied Phys.*, reference 23.)

[31] See reference 23.

[32] See section 11·10.

ary Eq. 17·49. Then the electrode strips representing the cathode and the anode are altered in shape until the potential along this insulating strip satisfies Eq. 17·50 to a sufficient accuracy. It is important to note that the current density j enters into the expression for the potential distribution, Eq. 17·50, only in a constant multiplying factor. Hence the electrode shapes obtained by the method outlined apply for all current densities (with space-charge saturation). It is only necessary to vary the voltage on the anode to obtain different currents, without danger of space-charge spreading.

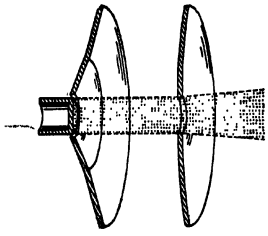


FIG. 17·30. Electrodes for Forming Parallel Space-Charge-Limited Beam. (Courtesy of *J. Applied Phys.*, reference 23.)

The electrode shapes found by Pierce for the case considered are shown in Fig. 17·30. The cathode is a flat disk cathode. The anode, similarly, should be a flat disk within the beam in order to satisfy the boundary conditions perfectly. Instead a hole has been bored in it to permit the passage of the beam. This causes the beam to diverge slightly, the lens action of the aperture being given by the formula for the refractive power $1/f = (\Phi_2' - \Phi_1')/(4\Phi)$, Φ_1' and Φ_2' being the potential gradients to the left and to the right of the aperture, respectively.

REFERENCES

1. O. SCHERZER, "Some defects of electron lenses," *Z. Physik*, Vol. 101, pp. 593–603, 1936.
2. R. REBSCH, "The theoretical resolving power of the electron microscope," *Ann. Physik*, Vol. 31, pp. 551–560, 1938.
3. W. GLASER, "A magnetic field free from spherical aberration," *Z. Physik*, Vol. 116, pp. 19–33, June 1940.
4. A. RECKNAGEL, "Spherical aberration in electron-optical image formation," *Z. Physik*, Vol. 117, pp. 67–73, December 1940.
5. R. REBSCH and W. SCHNEIDER, "Aperture defect of weak electron lenses," *Z. Physik*, Vol. 107, pp. 138–143, 1937.
6. O. SCHERZER, "Weak electric unipotential lens of least spherical aberration," *Z. Physik*, Vol. 101, pp. 23–26, 1936.
7. G. N. PLASS, "Electrostatic electron lenses with a minimum of spherical aberration," *J. Applied Phys.*, Vol. 13, pp. 49–55, January 1942; Vol. 13, p. 542, August 1942.
8. E. G. RAMBERG, "Variation of axial aberrations of electron lenses with lens strength," *J. Applied Phys.*, Vol. 13, pp. 582–594, September 1942.
9. J. DOSSE, "Optical data of strong electron lenses," *Z. Physik*, Vol. 117, pp. 722–753, 1941.
10. W. GLASER, "Short magnetic lens of least aperture defect," *Z. Physik*, Vol. 109, pp. 700–721, July 1938.

11. D. W. Epstein, "Electron-optical system of two cylinders as applied to cathode-ray tubes," *Proc. Inst. Radio Engrs.*, Vol. 24, pp. 1095–1139, August 1936.
12. E. Gundert, "Aperture defect of electrostatic cylinder lenses," *Z. Physik*, Vol. 112, pp. 689–690, June 1939.
13. E. Gundert, "The aperture defect of electrostatic electron lenses," *Die Telefunkenrohre*, No. 19–20, pp. 61–98, March 1941.
14. K. Spangenberg and L. M. Field, "Some simplified methods of determining the optical characteristics of electron lenses," *Proc. Inst. Radio Engrs.*, Vol. 30, pp. 138–144, 1942.
15. M. v. Ardenne, "Electrostatic high-voltage lens of short focal length," *Naturwissenschaften*, Vol. 27, pp. 614–615, September 1939.
16. E. Ruska, "Magnetic objective for the electron microscope," *Z. Physik.*, Vol. 89, pp. 90–128, May 1934.
17. H. Becker and A. Wallraff, "Spherical aberration of magnetic lenses," *Arch. Elektrotech.*, Vol. 32, pp. 664–675, October 1938.
18. W. Glaser, "Exact calculation of magnetic lenses with the field distribution $H = H_0/(1 + [z/a]^2)$," *Z. Physik*, pp. 285–315, 1941.
19. J. Dosse, "Exact calculation of magnetic lenses of asymmetric field distribution with $H = H_0/(1 + [z/a]^2)$," *Z. Physik*, Vol. 117, pp. 316–321, 1941.
20. H. Riedl, "Third-order aberrations of short weak purely electric unipotential electron lenses," *Z. Physik*, Vol. 107, pp. 210–216, 1937.
21. H. Voit, "Third-order electron-optical aberrations," *Z. Instrumentenk.*, Vol. 59, pp. 71–82, February 1939.
22. G. A. Morton and E. G. Ramberg, "Electron optics of an image tube," *Physics* Vol. 7, pp. 451–459, December 1936.
23. J. R. Pierce, "Rectilinear electron flow in beams," *J. Applied Phys.*, Vol. 11, pp. 548–555, August 1940.

CHAPTER 18

HIGH-VOLTAGE ELECTRON OPTICS. ION OPTICS

18·1. Magnetic Lenses. Up to this point it has been tacitly assumed that the electron velocities within the electron lenses were sufficiently small compared to the velocity of light to permit a neglect of the variation of the electron mass with its velocity. Under these circumstances all path equations could be derived from the variation principle

$$\delta \int \left\{ \phi^{1/2} - \left(\frac{e}{2mc^2} \right)^{1/2} A \cos \chi \right\} ds = 0 \qquad [18\cdot1]$$

ϕ denoting the electric potential, A the magnetic vector potential, and χ the angle between the vector A and the element of path ds. It has already been seen (Eq. 10·29) that, if the assumption of low electron velocities is dropped, the more generally valid variation principle

$$\delta \int \left\{ \left(\phi + \frac{e\phi^2}{2mc^2} \right)^{1/2} - \left(\frac{e}{2mc^2} \right)^{1/2} A \cos \chi \right\} ds = 0 \qquad [18\cdot2]$$

applies. If the fields considered are purely magnetic (ϕ = const), this leads to a relatively minor change. The potential ϕ must merely be replaced everywhere by the somewhat larger constant

$$\phi + \frac{e\phi^2}{2mc^2} = \phi(1 + 0.978\cdot10^{-6}\phi)$$

For example, the refractive power of a short magnetic lens becomes

$$\frac{1}{f} = \frac{0.0220}{\phi(1 + 0.978\cdot10^{-6}\phi)} \int H^2 \, dz \qquad [18\cdot3]$$

It is thus seen that for extremely high voltages (many million volts) the refractive power of a magnetic lens decreases with the reciprocal of the square of the potential rather than with the reciprocal of the first power. Figure 18·1 shows the change in refractive power $1/f$ and in the spherical aberration (Cf) for fixed aperture angle for the two-cylinder magnetic lens, System C in Fig. 17·3. Here for each curve the magnetic field is kept constant, identical markers indicating corresponding curves for the refractive power and the spherical aberration. A scale indicating the relative value of the least resolvable separation $d_{\min}$ ($d_{\min} \propto [Cf]^{1/4}$) for the lens employed as electron-microscope objective is also given.

Here the abscissa v represents the true applied voltages, whereas in the graphs of Chapter 17 (in particular Fig. 17·5) V is equivalent to the quantity $\phi + 0.978 \cdot 10^{-6}\phi^2$.

Equation 18·3 indicates that the change in focal length due to the relativistic variation of the mass of the electron is directly proportional

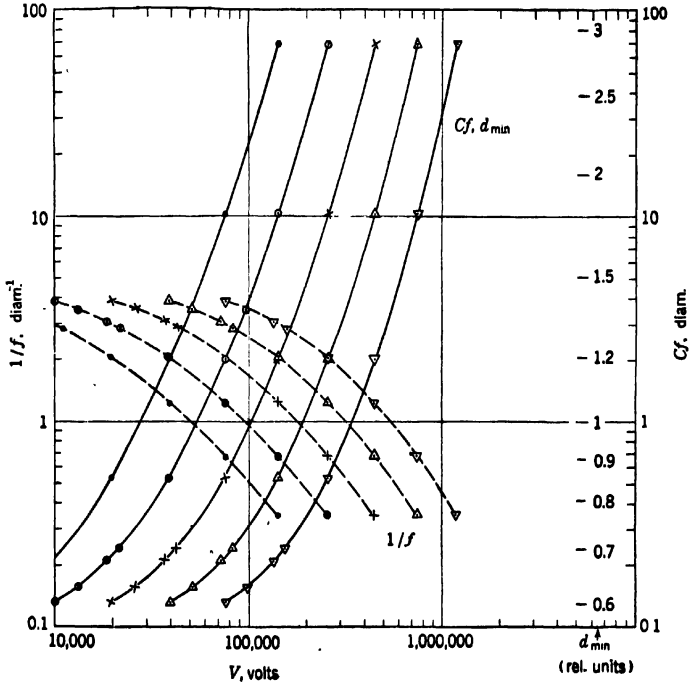


FIG. 18·1. Variation of Refractive Power and Spherical Aberration of a Magnetic Lens (System C, Fig. 17·3) with Applied Voltage. (Courtesy of *J. Applied Phys.*, Vol. 13, pp. 582–594, 1942.)

to the applied voltage, being 1 per cent for 10,000 volts, 2 per cent for 20,000 volts, 5 per cent for 50,000 volts, etc.

18·2. Electrostatic Lenses. For electrostatic fields the effect of the variation of mass is more complex. Thus, it is evident from Eq. 18·2 (with $A = 0$) that, taking account of the correcting term, the integrand no longer consists of a product of $\phi^{1/2}$ with an inconsequential constant multiplying factor. If the electron velocities are sufficiently high, changes in the applied voltage will change the shape of the electron paths even if all potentials maintain the same ratio.

Consider electron paths in a meridional plane and put, for convenience,

$$\frac{e}{2mc^2} = 0.978 \cdot 10^{-6} \text{ volt}^{-1} = a \qquad [18\cdot4]$$

Then the variation principle takes the form

$$\delta \int_{z_o}^{z_i} \{(\phi + a\phi^2)(1 + r'^2)\}^{1/2}\, dz = \delta \int_{z_o}^{z_i} F(r,z;r')\, dz = 0$$

and its Euler equation

$$\frac{d}{dz}\frac{\partial F}{\partial r'} - \frac{\partial F}{\partial r} = 0$$

becomes

$$r'' = \frac{1 + r'^2}{2\phi}\left(\frac{1 + 2a\phi}{1 + a\phi}\right)\left(\frac{\partial \phi}{\partial r} - r'\frac{\partial \phi}{\partial z}\right) \qquad [18\cdot5]$$

For paraxial electrons this may be written, introducing the expansion Eq. 11·75 for ϕ and retaining only first-order terms in r and r':

$$r'' + \frac{r'\Phi'}{2\Phi}\left(\frac{1 + 2a\Phi}{1 + a\Phi}\right) + \frac{r\Phi''}{4\Phi}\left(\frac{1 + 2a\Phi}{1 + a\Phi}\right) = 0 \qquad [18\cdot6]$$

Thus the effect of the variation of the mass with velocity on the paraxial equation is to replace the axial potential Φ by the smaller quantity $\Phi(1 + a\Phi)/(1 + 2a\Phi)$. In terms of the dependent variable $R = r\Phi^{1/4}$ the paraxial equation is, finally,

$$R'' + \frac{3}{16}\left(\frac{\Phi'}{\Phi}\right)^2 R = -\frac{a\Phi}{1 + a\Phi} R\left(\frac{R'\Phi'}{2R\Phi} + \frac{\Phi''}{4\Phi} - \frac{\Phi'^2}{8\Phi^2}\right) = S(z) \qquad [18\cdot7]$$

This equation may be solved by the standard methods repeatedly invoked in deriving the expressions for the aberrations in Chapter 16. Thus, if R_α and R_γ are two independent solutions of the reduced equation (Eq. 18·7 with S set equal to zero) satisfying the initial conditions

$$\begin{aligned} R_\alpha(z_o) &= 0 \qquad R'_\alpha(z_o) = R'_o \\ R_\gamma(z_o) &= 1 \end{aligned} \qquad [18\cdot8]$$

the solution of Eq. 18·7 for a ray leaving the object plane $z = z_o$ with the initial values $R(z_o) = 0$, $R'(z_o) = R'_o$ becomes[1]

$$\begin{aligned} R = R_\alpha &\left\{1 + \frac{1}{R'_o}\int_{z_o}^{z} R_\gamma(z')S(z')\, dz'\right\} \\ &- \frac{R_\gamma}{R'_o}\int_{z_o}^{z} R_\alpha(z')S(z')\, dz' \end{aligned} \qquad [18\cdot9]$$

[1] See Adams, reference 1, Formula 8.410.

In particular, for $z = z_i$, that is, in the image plane for $a\Phi \to 0$,

$$\Delta R_i = R_i = -\frac{M}{R_o'}\left(\frac{\Phi_i}{\Phi_o}\right)^{1/4}\int_{z_o}^{z_i} R_a(z')S(z')\,dz' \qquad [18\cdot10]$$

where M denotes the magnification of the image for $a\Phi \to 0$.

This solution is exact. If $a\Phi \ll 1$, the corresponding quantity

$$\Delta r_i = -\frac{M}{r_o'\Phi_o^{1/2}}\int_{z_o}^{z_i} R_a(z')S(z')\,dz' \qquad [18\cdot11]$$

may be regarded as an aberration, which can appropriately be referred to as the " axial relativistic aberration." It represents the spreading, due to the variation of the mass of the electron with velocity, of the central image point as the voltage on the electrodes of the electrostatic electron lens considered is increased from zero to the maximum value employed, keeping the voltage ratios of all electrodes constant. With $a\Phi \ll 1$ it is legitimate to replace R in $S(z)$ by R_a, the solution of the reduced equation, and Eq. 18·11 becomes, dropping the subscript α for convenience,

$$\begin{aligned}\Delta r_i &= a\Phi_A\cdot M\cdot\Phi_o^{-1/4}\cdot R_a\frac{R_a}{R_o'}\int_{z_o}^{z_i}\frac{R^2}{R_a^2}\left(\frac{R'\Phi'}{2R\Phi_A}+\frac{\Phi''}{4\Phi_A}-\frac{\Phi'^2}{8\Phi\Phi_A}\right)dz \\ &= -\,a\Phi_A\cdot\frac{M}{8}\cdot\Phi_o^{-1/4}\cdot R_a\frac{R_a}{R_o'}\int_{z_o}^{z_i}\frac{R^2}{R_a^2}\frac{\Phi'^2}{\Phi\Phi_A}dz\end{aligned} \qquad [18\cdot12]$$

Here Φ_A is the anode voltage and R_a is the value of R_α at some arbitrary " aperture plane " in image space. Since M and $r_i'\cdot R_o'$ are always of the same sign, the shift in the image plane $\Delta z_i = -\Delta r_i/r_i'$ is always positive, that is, the lens decreases in refractive power with increasing voltage.

For very large image distances the relativistic aberration on the axis takes on a particularly simple form:

$$\frac{\Delta r_i}{M} = -\frac{a\Phi_A f_i}{8}f_o\theta\int_{z_o}^{z_i}\frac{R^2}{R_a^2}\frac{\Phi'^2}{\Phi\Phi_A}dz = -K\Phi_A f_o\theta \qquad [18\cdot13]$$

θ signifying the aperture angle of the ray and K being given by

$$K = 1.223\cdot10^{-7}f_i\int_{z_o}^{z_i}\frac{R^2}{R_a^2}\frac{\Phi'^2}{\Phi\Phi_A}dz \text{ volt}^{-1} \qquad [18\cdot14]$$

Here R_a is the axial separation of the ray in image space after having left the lens. For weak electrostatic unipotential lenses K is a universal constant:

$$K = 0.652\cdot10^{-6} \text{ volt}^{-1}$$

since the integral in Eq. 18·14 represents, but for a constant factor, the refractive power of the lens.

Quite generally, substituting the integral expression for f_o, Eq. 18·14 may be written for a unipotential electrostatic lens:

$$K = 0.652 \cdot 10^{-6} \frac{\int_{z_o}^{z_i} \frac{\Phi}{\Phi_A} \frac{R^2}{R_a^2} \frac{\Phi'^2}{\Phi^2} dz}{\int_{z_o}^{z_i} \frac{R}{R_a} \frac{\Phi'^2}{\Phi^2} dz}$$

Since $R/R_a \leq 1$, K is numerically smaller than $0.652 \cdot 10^{-6}$ volt^{-1} if, in addition, $\Phi/\Phi_A \leq 1$. It follows that the relativistic aberration of a

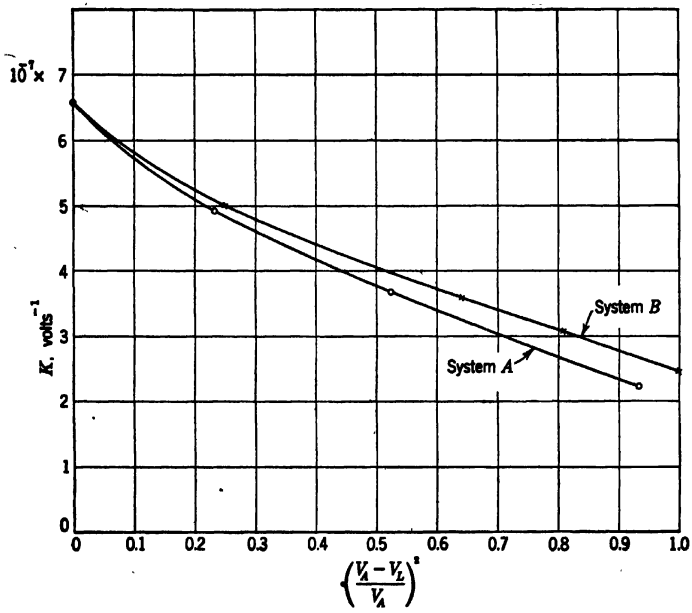


FIG. 18·2. Variation of Axial Relativistic Aberration of Two Unipotential Lenses (Systems A and B, Fig. 17·3) with Applied Voltage. (Courtesy of *J. Applied Phys.*, Vol. 13, pp. 582–594, 1942.)

strong electrostatic lens with negative center electrode is invariably smaller than that of a weak lens (for fixed height of incidence of the rays on the lens). Figure 18·2, showing a plot of K as function of lens strength for the two unipotential lenses Systems A and B (Fig. 17·3) with negative center electrodes, indicates that, in the range covered, the relativistic aberration drops by a factor equal to nearly 3. For electrostatic lenses with positive center electrodes the converse behavior must be

expected, since the factor $\Phi/\Phi_A \geq 1$ is likely to dominate the factor $R/R_a \leq 1$.

In analogy with chromatic aberration, there exists, in addition to axial relativistic aberration, a relativistic change in magnification which is dependent on the position of the aperture plane and affects the sharpness of the outer portions of an image produced by an electrostatic lens operated with fluctuating overall voltage.

An idea of the magnitude of the relativistic aberration may be gained by considering a specific example. Let the operating voltage of an electrostatic microscope objective such as System B with the center electrode connected to the cathode ($V_L = 0$; $K = 2.46 \cdot 10^{-7}$ volt^{-1}) fluctuate about its mean value with an amplitude of A volts. If the effective aperture angle is assumed to be $3 \cdot 10^{-3}$ radian and the focal length to be 1 centimeter, what is the value of A which will lead to a circle of confusion due to relativistic aberration with a diameter equal to $20M$ A.U.? The answer is given by

$$2 \cdot 10^{-7} = 2 \cdot 2.46 \cdot 10^{-7} \cdot A \cdot 3 \cdot 10^{-3} \qquad A = 135 \text{ volts}$$

For a circle of confusion $200 \cdot M$ A.U. in diameter the permissible amplitude of the fluctuation would be ten times as great. As long as $a\Phi \ll 1$ the permissible amplitude remains independent of the magnitude of the operating voltage. The permissible percentage fluctuation, however, is inversely proportional to the applied overall voltage.

In a number of applications of electron optics the relativistic effect cannot be regarded as an aberration, $a\Phi$ being of the order of, or even substantially greater than, unity. Hansen and Webster[2] have treated this case for a set of coaxial tubes at different potentials, assuming that the energy change of the electron in passing from one tube to the next is small compared to its initial energy. With the aid of Eq. 18·6, it is possible to derive the Gaussian focusing properties of electrostatic lenses at arbitrarily high voltages without any restrictions of generality. Again, the introduction of a dependent variable which eliminates the first-derivative term from the differential equation is of advantage. Substitute in Eq. 18·6 $r = P/g(z)$. The requirement that the term with P' in the resulting equation should vanish leads to a differential equation involving $g(z)$ and Φ, which is solved by

$$g(z) = \Phi^{1/4}(1 + a\Phi)^{1/4} \qquad [18\cdot15]$$

Thus, with $P = r\Phi^{1/4}(1 + a\Phi)^{1/4}$,

$$P'' = -\tau(z) \cdot P \qquad \tau(z) = \frac{3\Phi'^2}{16\Phi^2}\left\{1 - \frac{2}{3}\frac{a\Phi}{1 + a\Phi} + \frac{a^2\Phi^2}{(1 + a\Phi)^2}\right\} \qquad [18\cdot16]$$

[2] See reference 2.

A single integration of this equation for $P_i' = 0$ leads to the formula for the focal length of the short electrostatic lens at arbitrary voltage:

$$\frac{1}{f_o} = -\frac{r_o'}{r_i} = -\frac{P_o'}{P_i}\left(\frac{\Phi_i(1+a\Phi_i)}{\Phi_o(1+a\Phi_o)}\right)^{1/4} = \left(\frac{\Phi_i(1+a\Phi_i)}{\Phi_o(1+a\Phi_o)}\right)^{1/4}\int \tau(z)\,dz \qquad [18\cdot17]$$

In particular, for $a\Phi \to \infty$, that is, extremely high potentials

$$\frac{1}{f_o} = \frac{1}{4}\left(\frac{\Phi_i}{\Phi_o}\right)^{1/2}\int_{z_o}^{z_i}\frac{\Phi'^2}{\Phi^2}\,dz \qquad [18\cdot18]$$

For a unipotential lens, in particular, this formula differs from that valid for low voltages only by the factor ¼ replacing 3/16 — the refractive power at very high voltages is just 33⅓ per cent greater than at low voltages. An examination of the function $\tau(z)$ indicates that the refractive power decreases initially with increasing voltage. For a weak lens the minimum, equal to eight-ninths of the refractive power at low voltages, is reached for $a\Phi = 0.5$, that is, for an operating voltage of approximately 500,000 volts. Beyond this point, the refractive power increases once more, eventually approaching the value given by Eq. 18·18.

18·3. Compound Lenses at High Voltages; Aberrations. In the more general case of skew rays and in the presence of a magnetic field the paraxial equation may be derived from the general variation principle, Eq. 18·2, in the same manner as Eq. 15·58 was derived from the restricted variation principle, Eq. 18·1. The equation becomes

$$r'' = -\frac{\Phi'}{2\Phi}\left(\frac{1+2a\Phi}{1+a\Phi}\right)r' - \left\{\frac{\Phi''}{4\Phi}\left(\frac{1+2a\Phi}{1+a\Phi}\right) + \frac{eH^2}{8mc^2(1+a\Phi)\Phi} - \frac{C^2}{r^4(1+a\Phi)\Phi}\right\}r \qquad [18\cdot19]$$

Here the constant C is a measure for the angular momentum of the electron about the optic axis and vanishes for rays crossing the axis at any point. It is related to the rate of change of the azimuth θ of the electron by the equation

$$C = r^2(\Phi + a\Phi^2)^{1/2}\theta' - r^2\left(\frac{eH^2}{8mc^2}\right) \qquad [18\cdot20]$$

The meridional plane in which the electron finds itself at any point

along its path, and hence the image rotation, is given by

$$\theta - \theta_o = \int_{z_o}^{z_i} \left\{ \frac{C}{r^2(\Phi + a\Phi^2)^{1/2}} + \left[\frac{eH^2}{8mc^2(\Phi + a\Phi^2)} \right]^{1/2} \right\} dz \qquad [18\cdot21]$$

Thus, a being given by Eq. 18·4, the first-order focusing properties — that is, focal points, principal planes, focal lengths, magnification and rotation of the image — can be determined by the numerical tracing of rays through given electrostatic and magnetic fields in just the same manner as in the limiting case of low voltages.

Furthermore, for purely magnetic lenses the third-order expressions for the geometrical aberrations remain unchanged in passing to high voltages, with the sole exception that Φ in the constant multiplying factors must be replaced by $\Phi + a\Phi^2$.

In the presence of electrostatic fields the situation becomes somewhat more complex. However, the derivation of the aberration expression in Eqs. 16·60 and 16·63 from the variation principle is such that these equations become valid for arbitrarily high voltages if ϕ is replaced by $\phi + a\phi^2$, that is, if in Eqs. 16·58 and 16·60 the following substitutions are made:

$$\begin{aligned} \Phi &\rightarrow \Phi(1 + a\Phi) \\ \Phi'' &\rightarrow \Phi''(1 + 2a\Phi) \\ \Phi^{IV} &\rightarrow \Phi^{IV}(1 + 2a\Phi) + 4a\Phi''^2 \end{aligned} \qquad [18\cdot22]$$

It is, furthermore, possible to eliminate from the resulting expressions the higher derivatives, so as to obtain formulas analogous to Eqs. 16·64.

The expressions for the chromatic aberrations of electron lenses are derived very readily from Eq. 18·19 in the same manner in which they are derived from its low-voltage counterpart in section 16·6. They take the following form:

$$\begin{aligned} \Delta x_i + i\Delta y_i = \{&(C_1 + iC_2)r_o + C_3 r_a\} \\ &\exp i \left\{ \theta_o + \int_{z_o}^{z_i} \left(\frac{eH^2}{8mc^2(\Phi + a\Phi^2)} \right)^{1/2} dz \right\} \end{aligned} \qquad [18\cdot23]$$

where

$$C_1 = -\frac{M\Delta\Phi}{(\Phi_o + a\Phi_o^2)^{1/2}} \int_{z_o}^{z_i} \left\{ \left(\frac{r_\gamma' r_\alpha \Phi'}{2} + \frac{r_\gamma r_\alpha \Phi''}{4} \right) \frac{1 + 2a\Phi + 2a^2\Phi^2}{(\Phi + a\Phi^2)^{3/2}} \right. \qquad [18\cdot24a]$$

$$\left. + r_\gamma r_\alpha \frac{1 + 2a\Phi}{(\Phi + a\Phi^2)^{3/2}} \frac{eH^2}{8mc^2} \right\} dz$$

$$C_2 = -M\Delta\Phi \int_{z_o}^{z_i} \frac{1 + 2a\Phi}{2(\Phi + a\Phi^2)} \left(\frac{eH^2}{8mc^2(\Phi + a\Phi^2)} \right)^{1/2} dz \qquad [18\cdot24b]$$

$$C_3 = -\frac{M\Delta\Phi}{r_{a0}(\Phi_0 + a\Phi_0^2)^{1/2}}\int_{z_0}^{z_i}\left\{\left(\frac{r_a r_a'\Phi'}{2} + \frac{r_a^2\Phi'}{4}\right)\frac{1 + 2a\Phi + 2a^2\Phi^2}{(\Phi + a\Phi^2)^{3/4}}\right.$$

$$\left. + r_a^2\frac{1 + 2a\Phi}{(\Phi + a\Phi^2)^{3/4}}\frac{eH^2}{8mc^2}\right\}dz \qquad [18{\cdot}24c]$$

For a purely magnetic lens with large magnification M there follows from this the inequality

$$|C_3| \leq \left|\frac{M\Delta\Phi}{\Phi}\left(\frac{1 + 2a\Phi}{1 + a\Phi}\right)\right| \qquad [18{\cdot}25]$$

the expression on the right representing at the same time the value of the magnitude of C_3 for a short magnetic lens. It is seen from this that the chromatic aberration of a magnetic lens increases with applied voltage, reaching, at very high voltages, a value twice as great as in the limiting case of low voltages.

For an electrostatic short lens with large M the value of C_3 becomes:

$$C_3 = -2M\frac{\Delta\Phi}{\Phi}\frac{3 + 8a\Phi + 6a^2\Phi^2 + 4a^3\Phi^3}{3 + 7a\Phi + 8a^2\Phi^2 + 4a^3\Phi^3} \qquad [18{\cdot}26]$$

This is the same at very low and very high voltages and differs but little from the same value at intermediate voltages, the fraction being throughout very near to unity. Thus at very high voltages the chromatic aberration (for the same height of incidence and for the same, high, magnification) is identical for all electrostatic and magnetic short lenses.

18·4. Ion Optics. In sections 2·5 and 2·6 a number of devices are described in which electric and magnetic fields are used to focus or deflect atomic ions in place of electrons. All the important differences between the optical properties of ions and electrons can be deduced from the greater mass and the different sign of the usual positive ions. The index of refraction of an arbitrary field for an ion of charge E and mass M is, quite generally,

$$n = \left(-\frac{2E\phi}{Mc^2} + \frac{E^2\phi^2}{M^2c^4}\right)^{1/2} + \frac{E}{Mc^2}A\cos\chi \qquad [18{\cdot}27]$$

where the zero of the potential ϕ is chosen so that $-E\phi$ represents the kinetic energy of the particle. A again is the magnetic vector potential and χ the angle between it and the direction of motion of the ion.

For the lightest of all the positive ions, with the largest specific charge E/M, that is, the proton, $E = e$ and $M = 1837m$. Thus, for a positive ion the accelerating voltage or kinetic energy must be at least 1837 times as great as for an electron if the relativistic effect, represented by the

second term under the root in Eq. 18·23, is to assume equal importance in both cases. The relativity correction (and the velocity) of a 10-million-volt proton corresponds approximately to that of a 5000-volt electron. Hence the second term under the root may be neglected in the great majority of cases. Under these circumstances the paths in a purely electric field are given by

$$\delta \int (-\phi)^{1/2}\, ds = 0$$

that is, are identical with those of electrons in a similar field with the sign of the potential reversed and the relativity correction neglected. Furthermore, particles with a charge of the same sign and with the same ratio of initial kinetic energy to charge and the same initial direction of motion describe the same paths in electric fields, regardless of their mass. This also follows from the absence of the charge and mass from the ray equation (for example, Eq. 12·9) and from the formulas derived from it, such as that for the refractive power of a short lens (Eq. 13·34).

In combined electric and magnetic fields the effect of the magnetic field as compared with that of the electric field is reduced in proportion of $(E/M)^{1/2}$. Thus the refractive power of a magnetic lens, being proportional to the square of the magnetic field, is smaller for protons as compared with electrons by the factor $m/M = 1/1837$. Accordingly, a magnetic field is useful for separating ions of different mass and separating ions from electrons, but is relatively ineffective for either deflecting or focusing ions of high velocity. For such purposes electrostatic deflecting fields and electrostatic lenses are preferably employed.

Three other properties, which are consequences of the greater mass of the ions, distinguish the behavior of the ions from that of the electrons.

First of all, the greater mass of the ions corresponds to a smaller velocity for a given kinetic energy and hence to a greater space-charge concentration j/v (j = current density, v = velocity) for a given voltage and current. Ion beam currents must be less by a factor $(m/M)^{1/2} \leq 1/43$ if space-charge effects are to be no larger than for electrons of similar kinetic energy.

Second, the same reduced velocity for given kinetic energy results in a longer time of interaction between the ion and the atoms composing the matter through which it passes, leading to a greater loss of energy per unit distance of matter traversed than for electrons of equal energy. Ions penetrate matter less readily than electrons.

Third, the wave length of ions is smaller, for fixed accelerating voltage, than that of electrons by the factor $(m/M)^{1/2} \leq 1/43$. This circumstance

points to the theoretical possibility of constructing ion microscopes with resolving power greater than that of electron microscopes.

18·5. Relativistic and Focusing Effects in Accelerating Devices. It has already been seen that, for fixed accelerating potential, relativistic effects become less important in proportion to the inverse mass of the accelerated charged particles. Accordingly, the operation of the *linear accelerator*[3] perfected by Sloan and Lawrence,[4] employing mercury ions with a mass $3.7 \cdot 10^5$ times as great as that of the electron, may be expected to be practically unaffected by the change of mass with velocity. Here (Fig. 18·3) the essential part of the accelerator consists of a succession of coaxial tubes of increasing length. An alternating potential is applied between the odd- and even-numbered tubes, the frequency being

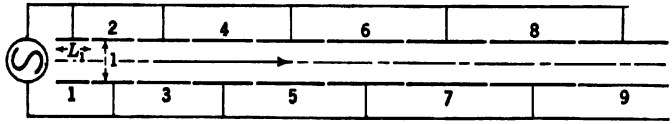


FIG. 18·3. Accelerating System of the Linear Accelerator According to Sloan and Lawrence (Schematic).

chosen so that an ion encountering a maximum accelerating field at one junction between tubes encounters accelerating fields in the same phase at succeeding junctions. In order that this may take place, the following relation must hold between tube length, L, velocity within the tube, v, and frequency, ν:

$$\frac{L}{v} = \frac{2k+1}{2\nu} \qquad k = 0, 1, 2, \cdots$$

It is customary to choose $k = 0$, so as to reduce the length of the tubes to their minimum value. In terms of the amplitude V of the alternating voltage and the initial kinetic energy eV_o of the ions, the length of the nth tube becomes:

$$L_n = \frac{1}{2\nu}\left(\frac{2e[V_o + nV]}{m}\right)^{1/2} \qquad [18\cdot28]$$

Thus the length of the tubes must increase approximately as the square root of their ordinal number. For the apparatus described by Sloan and Lawrence, consisting of thirty tubes with $L_1 = 1$ centimeter, $V_o = 10{,}000$ volts, and $V = 42{,}000$ volts, the total length of the set of tubes is 114 centimeters, the resonant frequency $\nu = 2.22 \cdot 10^7$ seconds^{-1}

[3] See section 2·5.
[4] See reference 3.

(13.5 meters), and the attainable energy of singly ionized mercury ions 1.3 Mev (million electron volts).

In order that a system of this type may have an appreciable output of high-velocity ions, two conditions must be fulfilled. First, the ions entering the system over an appreciable range of phase must be accelerated and pass through the system and, second, the action of the accelerating fields must be such as to cause the beam to converge rather than to increase its initial divergence.

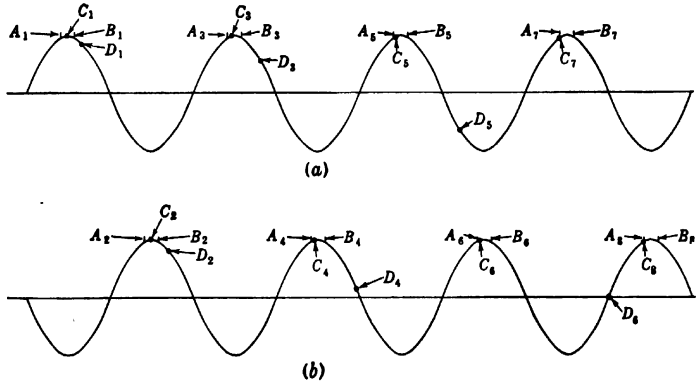


FIG. 18·4. Variation in Phase of Arrival of Ions at Junctions in the Linear Accelerator: (*a*) Voltage Wave on Odd Electrodes; (*b*) Voltage Wave on Even Electrodes.

The best arrangement for fulfilling the first condition adequately is to let the alternating voltage amplitude V be slightly greater than the value indicated by Eq. 18·28. Then the system is in exact resonance, that is, accelerates the ions in the same phase at all junctions, for the two phases A and B near the maximum of the voltage wave, as shown in Fig. 18·4. The two curves here represent the voltage waves on the sets of the odd and the even electrodes, respectively. It is seen that the system is stable, that is, ions will be continually accelerated, for initial phases between A and B and some distance to the left of A. Consider, for example, an ion entering the system in the phase C_1. Since it acquires slightly greater velocity than the ions at resonance, it will reach the next junction in a somewhat earlier phase C_2. This will continue until it has reached a phase well to the left of A, at which time its speed becomes less than that of the resonant ions (C_{11}). Accordingly, the phase of arrival at the junctions will move back toward and beyond A, subsequently oscillating about A, the ion being given an amount of

additional kinetic energy of the order of eV_A at each junction and, finally, leaving the system with an amount of energy approximately equal to $V_o + nV_A$. The same happens to an ion with an initial phase somewhat earlier than A. For an ion starting later than B, however, the phase of arrival D is progressively later until the ion is actually decelerated; thus such ions cannot contribute to the high-velocity ion beam.

Two factors influence the focusing: the lens properties of the fields at the junctions of the tubes (regarded as static lens fields) and the change in the strength of the field during passage of the ions. Since in the linear accelerator the continuously accelerated ions pass through the gaps close to the maxima of the voltage waves, the change in the field during passage is relatively small, and the first effect may be expected to predominate.

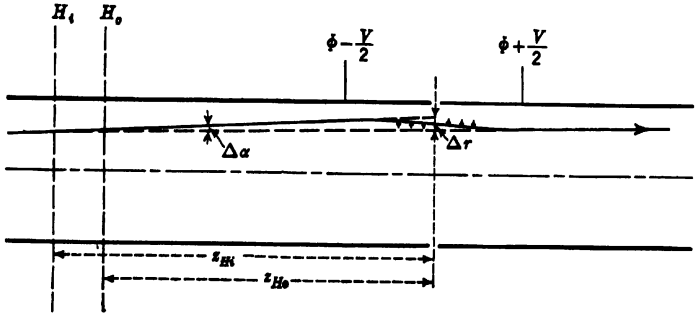


FIG. 18·5. Path of Ion in Field between Two Successive Tubes of the Linear Accelerator.

Figure 18·5 shows the path of an ion in the accelerating field between two cylinders. It is seen that the first part of the lens field exerts a converging force; the second, an equal diverging force. However, the diverging force, because of the greater velocity of the ion, is less effective, so that the net effect of the field may be described by a translation Δr toward the axis in the plane of symmetry of the gap and a decrease in slope $\Delta\alpha$ of the ray. These quantities can readily be deduced if the values of r and α for the incident ray, as well as the position of the principal planes and focal points, are known. Since, for all but the lowest values of n, the lenses may be regarded as weak, the thin-lens formulas of section 13·5 may be applied, setting the potential Φ within the lens equal to a constant. For a two-cylinder lens with

$$\Phi = \Phi_o + \left(\frac{V}{2}\right) \tanh 2.630z$$

using the cylinder diameter as unit of length, the positions of the principal planes and the focal lengths are given by

$$z_{H_o} = -\frac{1.52\Phi_o}{V} + 0.19 \qquad z_{H_i} = -\frac{1.52\Phi_o}{V} - 0.19$$

$$f_o = \left(\frac{\Phi_o - \frac{V}{2}}{\Phi_o + \frac{V}{2}}\right)^{1/4} f \qquad f_i = \left(\frac{\Phi_o + \frac{V}{2}}{\Phi_o - \frac{V}{2}}\right)^{1/4} f \qquad [18\cdot29]$$

$$\frac{1}{f} = 0.164\left(\frac{V}{\Phi_o}\right)^2$$

Here the separation between principal planes (0.38 diameter) may be disregarded relative to the distance of the principal planes from the plane of symmetry ($1.52\Phi_o/V$). Substitution of Eq. 18·29 in the lens equation, with Φ written for Φ_o,

$$\frac{\left(1 + \frac{V}{2\Phi}\right)^{1/4}}{b} - \frac{\left(1 - \frac{V}{2\Phi}\right)^{1/4}}{a} = \frac{\left(1 - \left[\frac{V}{2\Phi}\right]^2\right)^{1/4}}{f}$$

where a and b are the object and image distances measured from the principal plane, respectively, leads, with $\Phi/V = n - \frac{1}{2} \simeq n$, to the following expressions for $\Delta\alpha$ and Δr for large values of n:

$$\Delta\alpha = (r + z_H\alpha)\left[-\frac{\alpha}{r + z_H\alpha}\left(\frac{V}{2\Phi}\right) - 0.164\left(\frac{V}{\Phi}\right)^2\right] \qquad [18\cdot30a]$$

$$= -\frac{\alpha}{4n} - \frac{0.164}{n^2}$$

$$\Delta r = -z_H\cdot\Delta\alpha = -0.38\alpha - \frac{r}{4n} \qquad [18\cdot30b]$$

By writing $\Delta\alpha = d\alpha/dn$ and $\Delta r = dr/dn - \alpha L_1 n^{1/4}$, these equations become differential equations determining the variation of the slope and separation of the ion rays from the axis with the order number of the tube which they are entering. The term $\alpha L_1 n^{1/4}$ represents the change in axial separation during the passage through the nth tube. Solving Eq. 18·30a for r gives

$$r = -0.38\left(16n^2\frac{d\alpha}{dn} + 4n\alpha\right) \qquad [18\cdot31]$$

and substituting this in Eq. 18·30*b* leads to a differential equation for α alone:

$$\frac{d^2\alpha}{dn^2} + \frac{5}{2n}\frac{d\alpha}{dn} + \alpha\left(\frac{1}{4n^2} + \frac{0.164L_1}{n^{3/2}}\right) = 0 \qquad [18\cdot32]$$

With $\alpha = \beta n^{-1/4}$ this can be changed into

$$\frac{d^2\beta}{dn^2} + \beta\left(\frac{0.164L_1}{n^{3/2}} - \frac{1}{16n^2}\right) = 0 \qquad [18\cdot33]$$

Equation 18·33 is solved approximately, for $0.164L_1n^{-3/2} - 0.0625n^{-2} \gg 0$, by

$$\beta = \text{const}\left(\frac{0.164L_1}{n^{3/2}} - \frac{1}{16n^2}\right)^{-1/4} \sin\left\{\int\left(\frac{0.164L_1}{n^{3/2}} - \frac{1}{16n^2}\right)^{1/2} dn + \text{const}'\right\} \qquad [18\cdot34]$$

Thus, for very large n,

$$\alpha = C_1 n^{-1/8} \sin\,(1.62L_1^{1/2}n^{1/4} + \delta) \qquad [18\cdot35a]$$

$$r = C_2\{n^{3/8}\cos\,(1.62L_1^{1/2}n^{1/4} + \delta) - 1.54n^{1/8}L_1^{-1/2}\sin\,(1.62L_1^{1/2}n^{1/4} + \delta)\} \qquad [18\cdot35b]$$

where C_1, C_2, and δ are integration constants, depending on the initial slope and separation from the axis of the ion ray. Thus, for large n, the ion ray executes oscillations of slowly increasing amplitude and wave length. This behavior holds even for small n if the tube length, throughout, is large compared with the tube diameter. The increasing amplitude of the oscillations sets a limit to the number of accelerations which can be obtained without too greatly reducing the output of high-velocity ions.

The change of field during the transit of the ions through a lens will contribute to the focusing action if the field is decreasing in strength; for in this case the radial force is reduced in the diverging portion of the lens as compared with that in the converging part. Thus the change of field will enhance the focusing action for ions in a phase near B (Fig. 18·4) and reduce it for ions in a phase close to A. Since it has been seen that the phase of the ions will in general oscillate about A, the net effect of the changing field on the focusing action of the accelerating fields is unfavorable.

The *cyclotron* (Fig. 18·6) may, in a sense, be regarded as a linear accelerator wound up into a spiral, the path distance between successive accelerations being adjusted automatically by a uniform magnetic field normal to the ion paths and to the electric accelerating fields. Ions formed by the bombardment of the gas in the space between the two *dees* or *duants* forming the electrodes are drawn by the alternating field

between the dees into their interior. Here they describe semicircular paths under the influence of the magnetic field, the radius being given by

$$\frac{1}{R} = \left\{\frac{EH^2}{2Mc^2\left(\phi + \frac{E\phi^2}{2Mc^2}\right)}\right\}^{\frac{1}{2}} \cong \left(\frac{EH^2}{2Mc^2\phi}\right)^{\frac{1}{2}} \qquad [18\cdot36]$$

and the time required to traverse a semicircular path by

$$\frac{\tau}{2} = \frac{\pi R}{v} = \frac{\pi Mc}{EH}\left(1 + \frac{E\phi}{Mc^2}\right) \cong \frac{\pi Mc}{EH} \qquad [18\cdot37]$$

This time (to the nonrelativistic approximation) is quite independent of the velocity of the particle. Thus, if the frequency of the electrostatic field between the dees is made

$$\nu = \frac{EH}{2\pi Mc}$$

and the magnetic field is perfectly homogeneous, an ion in the median plane, which is accelerated in a certain phase at a particular crossing, will continue to be accelerated in the identical phase at all subsequent crossings. Thus the cyclotron, unlike the linear accelerator, is " resonant " for all phases of entrance of the ions into the dees.

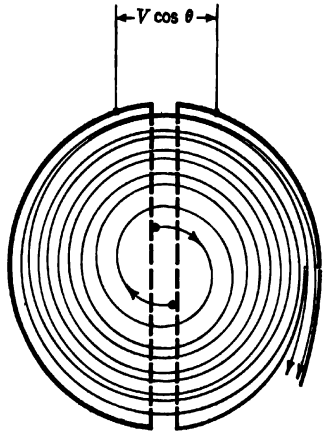


FIG. 18·6. Ion Paths in Cyclotron Provided with Two Ion Sources so as to Utilize the Full Cycle.

Similar electrostatic focusing effects control the vertical motion and radial motion, respectively, in the two cases. The space between the dees is, in effect, an electrostatic cylinder lens with properties similar to the lens at the junction of two tubes at different potentials. Here, also, the focusing effect of the lenses considered as static is superposed on the effect of the changing field. However, owing to the relatively large separation of the dees and the utilization of all initial phases, the effect of the changing fields is normally the more important one.[5] Since the changing field contributes to the convergence only for decreasing field strength (positive phase angle θ), the ion beam converges under the influence of the electric fields only for $0 < \theta < \pi/2$ and for $\pi < \theta < 3\pi/2$, the ions initially entering opposite dees in the two ranges of phase. Here,

[5] See reference 4.

according to Rose,[5] the vertical displacement is then approximately

$$z = \text{const}\left(\frac{\Phi}{\sin\theta}\right)^{1/4} \sin\left\{\int\left(\left[\frac{\pi V_o}{\Phi}\right]\sin\theta\right)^{1/2} dn + \delta\right\} \qquad [18\cdot38]$$
$$= \text{const}\; n^{1/4} \sin\left[2(\pi n)^{1/2} + \delta\right]$$

the second expression applying only for perfect resonance. This represents again an oscillation of gradually increasing amplitude and increasing period.

With this type of focusing alone, the cyclotron would be quite inefficient, the greater part of the ion beam eventually striking the top and bottom of the dees. As shown in Fig. 18·7, however, the

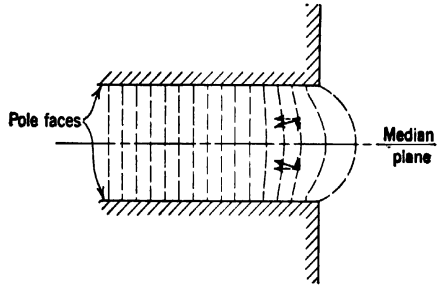


FIG. 18·7. Focusing Effect of the Magnetic Field Near the Edge of the Pole Faces. (Courtesy of *Phys. Rev.*, reference 4.)

bulging outwards of the magnetic field lines corresponding to the falling off of the field near the edge of the pole pieces exerts a force on the ions, pushing them back toward the median plane. This force causes the ions to oscillate about the median plane with decreasing amplitude and is much more effective in the focusing of the ions in the outer portions of their path than the lens action of the accelerating fields. This may be seen in detail by applying the equation of motion for the z-component of the ion velocity and making use of the fact that the radial magnetic field component is equal to a first approximation to the product of the distance z from the median plane and dH/dr — the rate of change of the magnetic field in the median plane with the radial distance.[6] In the nonrelativistic approximation

$$\frac{d^2z}{dn^2} = \left(\frac{\pi r}{v}\right)^2 \frac{d^2z}{dt^2} = \left(\frac{\pi r}{v}\right)^2 \frac{Eh_r v}{Mc}$$
$$= \left(\frac{\pi r}{v}\right)^2 \frac{Ez}{Mc} v \frac{dH}{dr} = \left(\pi^2 \frac{r}{H}\frac{dH}{dr}\right) z \qquad [18\cdot39]$$

[6] See section 15·5.

If the same procedure is applied for solving this equation as for obtaining Eq. 18·35*a*, it is seen that where the effect of magnetic field focusing predominates ($-\pi^2(r/H)dH/dr \gg \pi V \sin\theta/\Phi$) the vertical motion has a decreasing amplitude proportional to $\{-(r/H)dH/dr\}^{-\frac{1}{4}}$ and a decreasing period, the phase of the oscillation being given by

$$\int \pi\left(-\frac{r}{H}\frac{dH}{dr}\right)^{\frac{1}{2}} dn + \text{const}$$

Figure 18·8 shows the effects both of the accelerating fields alone and of the combined electric and magnetic focusing (dotted curve) on the paths of an ion as determined by Wilson.[7]

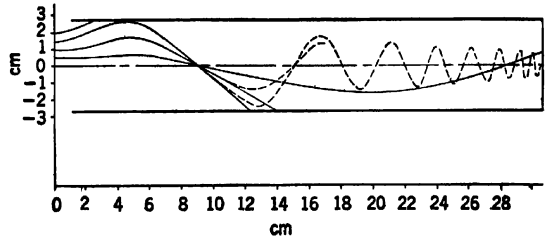


FIG. 18·8. Vertical Motion of the Ions in the Cyclotron as Determined by the Accelerating Fields Alone (———) and as Determined by the Combined Electric and Magnetic Focusing Action (- - - -). (Courtesy of *Phys. Rev.*, reference 5.)

Unfortunately the decrease of the magnetic field required for good focusing results in an imperfect fulfilment of the resonance condition (Eq. 18·37). This is made even more serious by the fact that the relativistic correction requires actually a field increasing toward the periphery for resonance. If the fulfilment of the resonance condition is relaxed in order to keep the ion beam convergent, the ions, after a certain number of revolutions, will be retarded in phase to such a degree that they will meet a decelerating field in place of an accelerating field between the dees and will hence spiral back toward the center of the cyclotron chamber. At best, ions will move outward and increase in energy only until they have been retarded in phase by π. From the requirement that the lag in phase angle does not exceed π and that, at the same time, the focusing condition is fulfilled throughout, Bethe and Rose[8] were able to establish an upper limit for the energy of ions emitted in appreciable quantities by the cyclotron. This maximum energy, which turns out to be proportional to the square root of both the mass and the charge of the particles, was found to be about 15 Mev for protons, 21 Mev for deuterons, and

[7] See reference 5.

[8] See Rose, reference 4, and Bethe and Rose, reference 6.

42 Mev for alpha particles. Appropriate ways of forming optimum magnetic fields by the insertion of thin iron shims between the pole pieces have been indicated by Rose.[9]

Though there thus appears to be a definite limit to the ion energy obtainable with a cyclotron using a radially symmetric field, Thomas[10] was able to show that this difficulty may be circumvented by giving the magnetic field an angular variation with a period $\pi/2$. Thus, suppose that the field variation chosen results in the closed path shown in Fig. 18·9, the effect of the electrostatic field between the dees being omitted. The curvature of the path, and hence the distance from the center, is greatest at the points A, B, C, and D, for which the magnetic field is a maximum. In the region AE both the magnetic field and the radial distance are decreasing. Accordingly, the field lines must here curve backwards toward A, and the velocity component toward the center must cooperate with the field component toward A to exert a force $F_z = E{\cdot}v_r{\cdot}h$ directed toward the median plane on the ions. Similarly, in section EB, where both field and radial distance are increasing, the field lines bend forward toward B and the corresponding force is again directed toward the median plane. Thus, throughout the cycle, the angular field variation produces a focusing effect. Thomas showed, in detail, that it was possible to obtain by this means a net focusing effect even with exact fulfilment of the resonance condition.

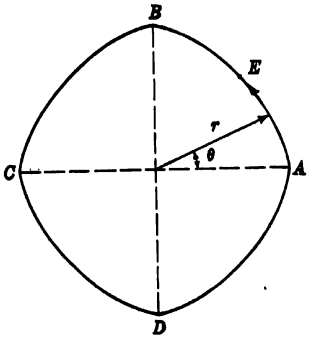


FIG. 18·9. Effect of an Angular Field Variation on an Ion Path in the Median Plane. (Courtesy of *Phys. Rev.*, reference 8.)

In the third instrument here considered, the magnetic induction accelerator or betatron perfected by Kerst,[11] the particles accelerated are electrons and, consequently, relativistic effects can no longer be regarded as a mere correction.

The operation of the instrument is based on the fact that a changing magnetic field has associated with it an electric field normal to it, as demanded by Maxwell's equation

$$\frac{1}{c}\frac{\partial H}{\partial t} = -\text{curl } E \qquad [18\cdot40]$$

[9] See reference 7.
[10] See reference 8.
[11] See Kerst, references 9 and 11 and Kerst and Serber, reference 10.

which by Stokes's law is equivalent to

$$\oint E\,ds = -\frac{1}{c}\int\left(\frac{\partial H}{\partial t}\right)_n dS \qquad [18\cdot41]$$

ds being an element of the boundary of the surface S, over which the normal component of $\partial H/\partial t$ is integrated. For a radially symmetric magnetic field the electric field E at a distance r from the center of symmetry is thus given by

$$-E = \frac{1}{2\pi rc}\int\frac{\partial H}{\partial t}\,dS = \frac{1}{2\pi rc}\frac{\partial}{\partial t}\int H\,dS = \frac{r}{2c}\frac{\partial H_m}{\partial t} \qquad [18\cdot42]$$

H_m being the average magnetic field within the circle of radius r.

If $e\phi$ is the kinetic energy of the electron,

$$d\phi = -E\,ds = -Ev\,dt$$

$$= -\left\{\frac{E\left(\frac{2e\phi}{m}+\frac{e^2\phi^2}{m^2c^2}\right)^{1/2}}{\left(1+\frac{e\phi}{mc^2}\right)}\right\}dt$$

or, by integrating,

$$\left(\phi^2+\frac{2mc^2}{e}\phi\right)^{1/2} - \left(\phi_o^2+\frac{2mc^2}{e}\phi_o\right)^{1/2} = c\int_{t_o}^{t} E(t)\,dt \qquad [18\cdot43]$$

Suppose now that the electron moves continuously on the circle of radius R about the axis of symmetry. Then, by Eqs. 18·43 and 18·42:

$$\left(\phi^2+\frac{2mc^2}{e}\phi\right)^{1/2} - \left(\phi_o^2+\frac{2mc^2}{e}\phi_o\right)^{1/2} = \frac{R}{2}(H_m - H_{mo}) \qquad [18\cdot44]$$

A comparison of this equation with Eq. 18·36 for the radius of curvature of the path of a charged particle in a magnetic field shows that, in fact, if at all times $H_m = 2H$, that is, if the average field within the path of the electron is twice as large as the magnetic field at the electron, and if, furthermore, the relation between the initial energy of the electron and the initial magnetic field is given by

$$\frac{H_o}{\left(\phi_o^2+\frac{2mc^2}{e}\phi_o\right)^{1/2}} = \frac{1}{R} \qquad [18\cdot45]$$

the electron will continue to move on the identical circular orbit of radius R, its increase in energy being given by Eq. 18·44. The maximum energy of the electrons is determined by the maximum value of the mag-

netic field as well as by the value of R:

$$\phi_{max} = -\frac{mc^2}{e} + \left\{\frac{m^2c^4}{e^2} + \frac{R^2H^2_{m\,(max)}}{4}\right\}^{1/2} \qquad [18\cdot46]$$
$$= (2.62\cdot10^{11} + [150RH_{m\,(max)}]^2)^{1/2} - 5.1\cdot10^5$$

For example, for $H_{m\,(max)} = 7000$ gauss and $R = 19$ centimeters, $\phi_{max} = 20$ Mev.

It can readily be seen that the orbits are stable. For the region outside the orbit $H_m/H < 2$, so that the ratio of the accelerating force along a tangent to the orbit to the centripetal force is too low to keep the electron moving in a circle about the axis of symmetry; thus an electron originally outside will spiral inward toward the equilibrium orbit. Similarly, for the region inside the orbit $H_m/H > 2$, so that the tangential acceleration outbalances the centripetal acceleration, resulting in paths spiraling outward toward the equilibrium orbit. These conditions are maintained even if the magnetic field at the equilibrium orbit is decreasing radially, provided that the rate of decrease is not stronger than $dH/dr = -H/r$. Hence magnetic focusing, resulting in damped oscillations of the electrons in a vertical direction, parallel to the axis of symmetry, is consistent with simultaneous fulfilment of the stability requirements.

If the electrons start from a position outside the equilibrium orbit with a direction of motion inclined to the tangential direction, as well as if they have either too high or too low initial energies to continue on a circle about the center of symmetry, they execute damped oscillations about the spiral toward the equilibrium orbit. These facts are utilized to get an appreciable number of electrons to clear the injector structure located inside the magnetic field but outside the equilibrium orbit. To withdraw the accelerated electrons from this orbit, the ratio H_m/H is suddenly decreased (by saturation of the central core) or increased (by increasing the central flux suddenly with the aid of auxiliary coils). The electrons consequently spiral inward to impinge on a suitably located target, or outward, bombarding the injector structure, respectively. Their presence is indicated by the generation of highly penetrating electromagnetic radiation (gamma rays).

REFERENCES

1. E. P. Adams, *Smithsonian Mathematical Formulae and Tables of Elliptic Functions*, Smithsonian Institution, Washington, 1922.
2. W. W. Hansen and D. L. Webster, "Electrostatic focusing at relativistic speeds," *Rev. Sci. Instruments*, Vol. 7, pp. 17–23, January 1936.
3. D. H. Sloan and E. O. Lawrence, "Production of heavy high-speed ions without the use of high voltages," *Phys. Rev.*, Vol. 38, pp. 2021–2032, 1931.

4. M. E. Rose, "Focusing and maximum energy of ions in the cyclotron," *Phys. Rev.*, Vol. 53, pp. 392–408, 1938.
5. R. R. Wilson, "Magnetic and electric focusing in the cyclotron," *Phys. Rev.*, Vol. 53, pp. 408–420, 1938.
6. H. A. Bethe and M. E. Rose, "The maximum energy obtainable from the cyclotron," *Phys. Rev.*, Vol. 52, pp. 1254–1255, 1937.
7. M. E. Rose, "Magnetic field corrections in the cyclotron," *Phys. Rev.*, Vol. 53, pp. 715–719, 1938.
8. L. H. Thomas, "The paths of ions in the cyclotron," *Phys. Rev.*, Vol. 54, pp. 580–598, 1938.
9. D. W. Kerst, "The acceleration of electrons by magnetic induction," *Phys. Rev.*, Vol. 60, pp. 47–53, 1941.
10. D. W. Kerst and R. Serber, "Electronic orbits in the induction accelerator," *Phys. Rev.*, Vol. 60, pp. 53–58, 1941.
11. D. W. Kerst, "New induction accelerator generating 20 Mev," *Phys. Rev.*, Vol. 61, pp. 93–94, 1942.

CHAPTER 19

IMAGE FORMATION IN THE ELECTRON MICROSCOPE

19·1. The Mechanism of Image Formation. An understanding of the process of image formation in the electron microscope is a prerequisite for a proper interpretation of the unfamiliar object detail revealed by

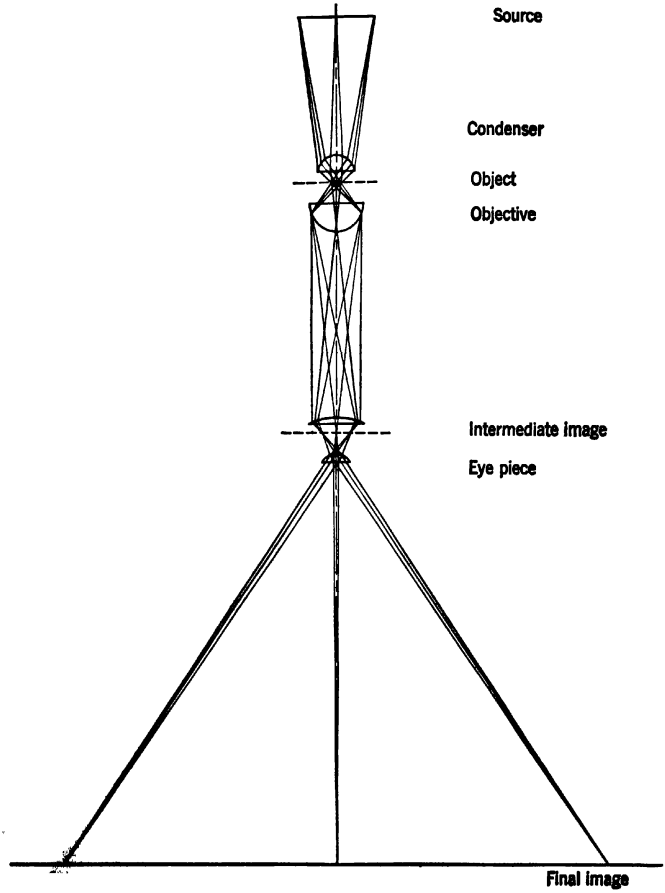


FIG. 19·1(a). Formation of the Image in a Light Microscope.

electron micrographs. In addition to this it may serve as a guide in the selection of suitable modes of operation of the instrument for the observation of different types of specimens.

Basically, the process resembles that taking place in the light microscope (Fig. 19·1*a*). A beam of electrons, the imaging medium, of controlled intensity and convergence is directed from the source onto the object. Different parts of the object, depending primarily on their mass thickness or " transparency," affect the electron stream incident on them differently, permitting a larger or smaller proportion thereof to enter the aperture of the objective. The objective focuses the electron pencils

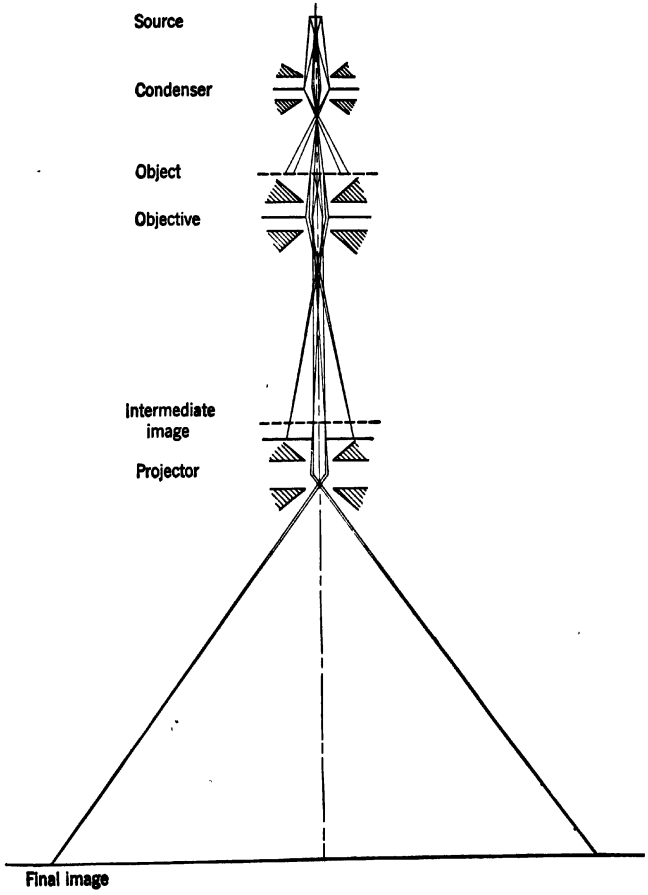


FIG. 19·1(*b*). Formation of the Image in an Electron Microscope.

proceeding from individual object points in a plane just above the projector lens, the intermediate image plane, forming there a first, magnified image of the object. The narrow pencils forming the central part of the intermediate image proceed unhindered through the projector lens, which causes a strong deflection of the off-axis pencils and forms a second, highly magnified image in the plane of the viewing screen or photographic plate. The analogy goes even farther: Replacing ray optics with wave optics, it may be shown that in either case the ultimate resolution of the instrument is determined — except for limits imposed by contrast — by the ratio of the wave length of the imaging medium and the effective numerical aperture of the objective.

There are, however, important differences to be considered. Thus the interaction of matter with electrons differs fundamentally from that with light. In addition to this, the uncorrected nature of the electron lenses normally employed necessitates the use of very small effective objective apertures.

19·2. Scattering and Absorption Processes in the Object. A multitude of interactions between light and matter determines the contrasts obtained in the image formed by a light microscope. Neutral absorption, spectrally selective absorption, specular reflection, refraction, scattering and diffraction in the object all play a role. In the electron microscope the situation is simpler. " Color " effects, or effects due to the selective action of different materials on electrons of different velocities, are entirely absent.[1] Similarly, refraction plays no role with fast electrons and absorption a very minor one. The contrasts observed are almost entirely due to differential scattering by different parts of the object, the degree of scattering being determined primarily by the mass density of the portion of the object considered. The larger the proportion of electrons which is scattered by a particular object element through an angle in excess of the effective aperture of the objective, the smaller is the intensity of the image of this same element.

[1] This is a consequence of the use of strictly monochromatic or *monokinetic* beams in the electron microscope, necessitated by the fact that electron lenses employed at present are chromatically uncorrected. However, certain selective absorption effects undoubtedly exist for fast electrons — in addition to the highly velocity-selective diffraction effects. The former effects are extremely small compared to the selective absorption effects for light in the visible part of the spectrum. Furthermore, the translation of differences in bombarding velocity into differences in color is, in principle, possible. For example, for visual observation a screen could be used consisting of several thin layers of fluorescent materials with different emission bands, the deeper layers being excited to a greater extent by the faster electrons. Similarly, for permanent record, " color plates " for electrons could be made by utilizing the property of thin films of matter of acting as intensifying screens for specific velocities as determined by the thickness of the films (see Baker, Ramberg, and Hillier, reference 1).

With crystalline objects diffraction by the lattice causes intensity variations in the image which bear no simple relationship to the distribution of mass density in the object. They will be discussed separately in section 19·9.

A high-speed electron may be scattered by an atom with or without appreciable loss of energy. Scattering which is not accompanied by such a loss is usually referred to as elastic scattering; scattering involving a decrease in kinetic energy, as inelastic scattering. In the latter case the lost energy is utilized to excite the atom or to eject one of the atomic electrons. If the electron is incident on solid substance, inelastic scattering may also result from electrostatic interaction with the "free" electrons of the material.

Precise quantitative data on the scattering processes can be obtained only by the application of quantum-mechanical methods.[2] A fairly adequate qualitative insight may be reached, however, with the aid of elementary classical considerations. Here elastic scattering is treated as the result of "collisions" of electrons with the (eventually shielded) nucleus of the atom, inelastic scattering as the consequence of electron-electron collisions. That a transfer of energy is involved in the latter case and none (or nearly none) in the former follow from the fulfilment of the laws of conservation of energy and momentum. If m is the mass of the incident electron, M that of the "struck" particle, be it nucleus or electron, and if $\dot{z}_o,\dot{r}_o = 0$; $\dot{Z}_o = 0$, $\dot{R}_o = 0$; $\dot{z}_1,\dot{r}_1$; $\dot{Z}_1,\dot{R}_1$ are their velocity components in their common plane before and after the collision, respectively (Fig. 19·2):

$$m\dot{z}_o = m\dot{z}_1 + M\dot{Z}_1 \qquad 0 = m\dot{r}_1 + M\dot{R}_1 \qquad [19\cdot1a]$$

$$m\dot{z}_o^2 = m(\dot{z}_1^2 + \dot{r}_1^2) + M(\dot{Z}_1^2 + \dot{R}_1^2) \qquad [19\cdot1b]$$

Assume first that $M \gg m$, that is, the case of a collision between an electron and a nucleus. The largest possible transfer of momentum (and hence of energy) to the originally stationary particle takes place if the electron reverses its direction of motion: $\dot{r}_1 = \dot{R}_1 = 0$,

$$m(\dot{z}_o - \dot{z}_1) = M\dot{Z}_1$$

$$m(\dot{z}_o^2 - \dot{z}_1^2) = M\dot{Z}_1^2$$

whence

$$M\dot{Z}_1^2 = m\dot{z}_o^2 \cdot \frac{4mM}{(M+m)^2} \cong \frac{m\dot{z}_o^2 \cdot 4m}{M} \qquad [19\cdot2]$$

[2] See Mott and Massey, reference 2, Marton and Schiff, reference 3, Williams, reference 4, and Goudsmit and Saunderson, reference 5.

Thus the maximum energy an electron can lose in a collision with a stationary proton, the lightest of all nuclei, is approximately 1/500 of its original energy. This corresponds to an angle of deflection of 180 degrees. For a smaller angle of deflection δ Eqs. 19·1 lead to a loss in the kinetic energy $e\Phi$ equal to

$$\Delta(e\Phi) = \frac{m\dot{z}_o^2}{2} \cdot \frac{2mM}{(M+m)^2} \left\{ 1 + \frac{m}{M} \sin^2 \delta - \cos \delta \left(1 - \left[\frac{m}{M} \right]^2 \sin^2 \delta \right)^{1/2} \right\} \qquad [19\cdot3]$$

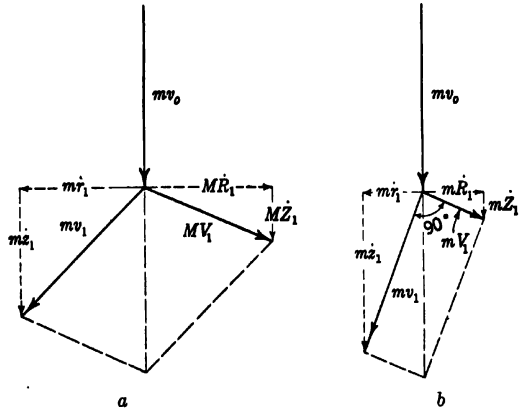


FIG. 19·2. Scattering of Electron by Nucleus (*a*) and by Electron (*b*). Momentum Balance.

For very large ratios M/m and relatively small angles of deflection δ, this becomes

$$\Delta(e\Phi) \cong \frac{m\dot{z}_o^2}{2} \frac{m}{M} \delta^2 \qquad [19\cdot3a]$$

For example, for an electron deflected by a hydrogen nucleus through an angle of 1 degree, the loss in energy is 1/6,000,000 of the original energy.

Equation 19·3 also indicates the loss in energy when an electron is scattered by another electron, which was originally at rest. With $M = m$:

$$\Delta(e\Phi) = \frac{m\dot{z}_o^2}{2} \sin^2 \delta \qquad [19\cdot3b]$$

The maximum deflection is in this case 90 degrees, the "struck" electron acquiring the full kinetic energy under these circumstances. Quite

generally, after collision, the velocity vectors of the two electrons form an angle of 90 degrees with each other, since (by Eq. 19·3*b*) they are equal in magnitude to their resultant, the initial velocity of the incident vector, multiplied by the cosine and by the sine of the angle of deflection. It is seen that for small angles the energy losses are larger than for nuclear collisions in the ratio M/m, which has a minimum value of 1837.

The equations above correlate the energy losses of electrons scattered by nuclei and initially stationary electrons with their angles of scattering, but yield no information with regard to the angular distribution of the scattered electrons. In order to determine the latter it is necessary to examine their paths in greater detail.

If it is assumed that the originally stationary particle of mass M has a charge Z^*e, the equations of motion of the two particles are

$$\ddot{z} = -\frac{Z^*e^2(z-Z)}{m[(r-R)^2+(z-Z)^2]^{3/2}}$$
$$\ddot{r} = -\frac{Z^*e^2(r-R)}{m[(r-R)^2+(z-Z)^2]^{3/2}} \qquad [19\cdot4a]$$

$$\ddot{Z} = -\frac{Z^*e^2(Z-z)}{M[(r-R)^2+(z-Z)^2]^{3/2}}$$
$$\ddot{R} = -\frac{Z^*e^2(R-r)}{M[(r-R)^2+(z-Z)^2]^{3/2}} \qquad [19\cdot4b]$$

It is convenient to translate these into the equations of motion of the center of gravity of the system (coordinates $\bar{Z}$, $\bar{R}$) and the equations for the motion of the incident electron relative to the originally stationary particle (coordinates ζ, ρ), respectively:

$$\bar{Z} = \frac{mz+MZ}{m+M} \qquad \bar{R} = \frac{mr+MR}{m+M}$$
$$\zeta = z - Z \qquad \rho = r - R \qquad [19\cdot5]$$

The center of gravity moves with uniform velocity in the original direction of the incident electron:

$$\bar{Z} = \frac{m}{m+M}\dot{z}_o(t-t_o) \qquad \bar{R} = \frac{m}{m+M}p \qquad [19\cdot6]$$

p being the initial distance of the particle of mass M from the asymptote to the path of the incident electron (Fig. 19·3). The initial position of the former particle has been chosen as origin.

On the other hand the equations

$$\ddot{\zeta} = -\left(\frac{m+M}{mM}\right) Z^* e^2 \frac{\zeta}{(\zeta^2+\rho^2)^{3/2}} \qquad [19\cdot7a]$$

$$\ddot{\rho} = -\left(\frac{m+M}{mM}\right) Z^* e^2 \frac{\rho}{(\zeta^2+\rho^2)^{3/2}} \qquad [19\cdot7b]$$

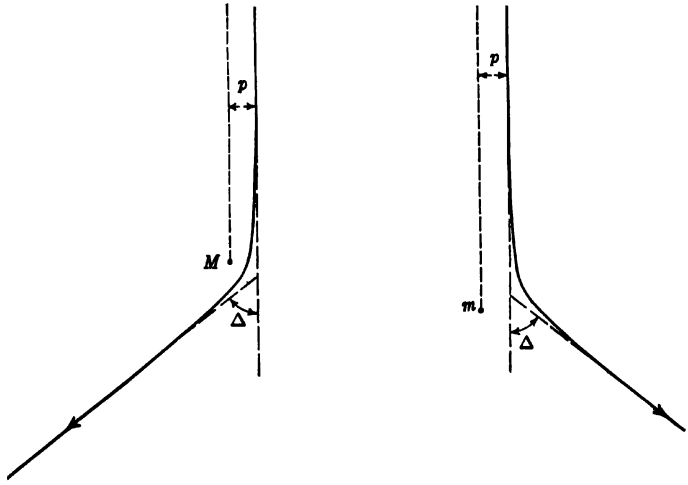


FIG. 19·3. Path of Electron Relative to Nucleus (*a*) and Other Electron (*b*) in Coordinate Systems Which Are Stationary Relative to the Latter Particle in Each Case.

describe the motion of a particle of mass $m^* = mM/(m+M)$ and charge $-e$ in the field of a stationary center of force with a charge Z^*e. The path of the electron relative to the originally stationary particle is thus a hyperbola with the particle, acting as a center of force, at one of the focal points — analogous, in every way, to the path of a comet about the sun.

If rectangular coordinates (Fig. 19·4) x, y are introduced with the x-axis as line of symmetry of the path and the origin placed at the intersections of the asymptotes to the hyperbola, the equation of the latter becomes

$$\frac{x^2}{a^2} - \frac{y^2}{p^2} = 1 \qquad [19\cdot8]$$

with a radius of curvature at the vertex $R_{\min} = p^2/a$. The value of a is

found from the law of equal areas

$$pv_o = \rho_{\min} v_{\max} \qquad [19\cdot9]$$

where v_o is the initial velocity of the incident electron, $\rho_{\min}$ its distance from the other particle as it passes through the vertex of the hyperbola,

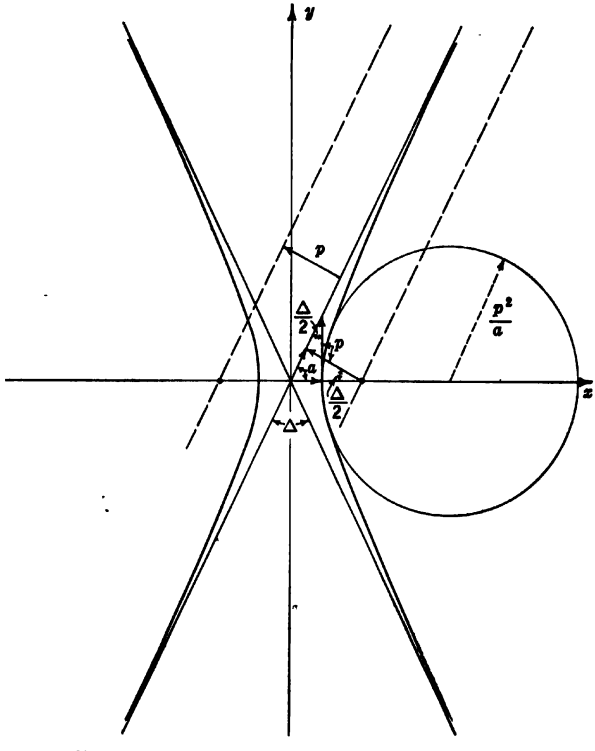


FIG. 19·4. Geometrical Properties of the Hyperbola.

and $v_{\max}$ its velocity at the same point. From the second law of motion applied at the vertex:

$$\frac{m^* v_{\max}^2}{R_{\min}} = \frac{Z^* e^2}{\rho_{\min}^2} \qquad [19\cdot10]$$

Equations 19·9 and 19·10 yield

$$a = \frac{Z^* e^2}{m^* v_o^2}$$

Finally, from the geometry of the hyperbola,

$$\tan\frac{\Delta}{2} = \frac{a}{p} = \frac{Z^*e(m+M)}{2\Phi pM} \qquad [19\cdot11]$$

where $e\Phi = mv_o^2/2$ is the initial kinetic energy of the incident electron. This yields the angle of deflection Δ in a coordinate system in which the originally stationary particle remains at rest. To find $\tan\delta = dr/dz$ it is necessary to return to the coordinates r, z of the fixed system.

Since in the moving system the collision is completely *elastic*, the center of force remaining fixed, the final velocity of the incident electron is equal to its original velocity v_o, so that the motion is given (for large t) by

$$\begin{aligned} \zeta &= \text{const} + v_o \cdot t \cdot \cos\Delta \\ \rho &= \text{const}' + v_o \cdot t \cdot \sin\Delta \end{aligned} \qquad [19\cdot12]$$

Furthermore, by Eq. 19·5,

$$z = \frac{\zeta M}{m+M} + \bar{Z} \qquad r = \frac{\rho M}{m+M} + \bar{R}$$

so that, by utilizing Eq. 19·6,

$$\tan\delta = \frac{dr}{dz} = \frac{v_o \dfrac{M}{m+M}\sin\Delta\, dt}{\left(v_o \dfrac{M}{m+M}\cos\Delta + \dfrac{mv_o}{m+M}\right)dt} = \frac{\sin\Delta}{\cos\Delta + \dfrac{m}{M}} \qquad [19\cdot13]$$

For $M \gg m$, $\delta \cong \Delta$, so that the angle of deflection is given directly by Eq. 19·11. On the other hand, for $M = m$ and $Z^* = -1$, that is, if the struck particle is an electron,

$$\tan\delta = \frac{\sin\Delta}{1+\cos\Delta} = \tan\frac{\Delta}{2} = -\frac{e}{p\Phi} = -\frac{1.44\cdot10^{-7}}{p\Phi} \qquad [19\cdot14]$$

The angles of scattering produced by a nucleus are thus, for equal closeness of approach, approximately Z^* times as large as those consequent on scattering by a second electron. The nuclei may be regarded as primarily responsible for the angular scattering, the electrons for the energy loss of the incident electrons.

When electrons pass through even a very thin layer of matter they are normally influenced by a number of atoms. However, deflections through a considerable angle may, on the average, be imparted to electrons in transit by a single atom only. If this is the case *single scattering* is said to take place for deflections by such or by larger angles, and the formulas given above are adequate (with limitations yet to be

specified) for calculating the angular distribution of the scattered electrons. The thicker the specimen, the larger the minimum angle for which single scattering applies. For relatively thick specimens and small scattering angles every electron undergoes so many deflections that they may be summed in accordance with the theory of errors, leading to a Gaussian distribution. This condition is termed *multiple scattering*. In the intermediate region, most often realized in electron-microscope specimens, *plural scattering* takes place. Each electron is deflected more than once on the average, yet not often enough to permit the application of the theory of errors to the determination of the angular distribution.

Consider, first, the case of single scattering. Let a uniform parallel stream of electrons impinge on a layer of substance containing Nx similar atoms of mass M and atomic number Z per square centimeter of superficial area, x being the thickness of the layer and $N = \rho N_A/A$ the number of atoms in a cubic centimeter of substance. A is the atomic weight, ρ the density, and N_A Avogadro's number (the number of atoms in a gram-atom). There will thus be Nx nuclei and NZx electrons per square centimeter. If it is assumed that these scatter the incident electrons without mutual interference, the electrons which will be scattered into a solid angle $2\pi \sin\delta d\delta$ comprising the conic shell between δ and $\delta + d\delta$ are incident on a fraction of the total area equal to

$$Nx \cdot 2\pi p_n dp_n + NZx \cdot 2\pi p_e dp_e \qquad [19\cdot15]$$

where p_n and p_e are given by Eqs. 19·11 and 19·14, respectively:

$$p_n = \frac{Ze}{2\Phi \tan\dfrac{\delta}{2}} \qquad p_e = \frac{e}{\Phi \tan\delta} \qquad [19\cdot16]$$

Substitution of these expressions in Eq. 19·15 leads to the following angular distribution of the scattered electrons:

$$I(\delta)d\Omega = \frac{Z^2e^2Nx}{16\Phi^2}\left\{\frac{1}{\sin^4\dfrac{\delta}{2}} + \frac{16}{Z}\frac{\cos\delta}{\sin^4\delta}\right\}\sin\delta d\delta\, d\phi \qquad [19\cdot17]$$

Here $d\Omega$ is an element of solid angle with the inclination δ and the azimuth ϕ. For small angles δ, the distribution becomes

$$I(\delta) = \frac{Z^2e^2Nx}{\Phi^2\delta^4}\left(1 + \frac{1}{Z}\right) \qquad [19\cdot17a]$$

Thus in spite of the greater number of electrons in the scattering layer, the number of incident electrons scattered by them into a given solid angle is only $1/Z$ times as large as that scattered by the nuclei.

Equations 19·17 and 19·17*a* are correct only for relatively large scattering angles and high-speed electrons, since the shielding of the nuclear charge by the surrounding electron cloud as well as the binding of the atomic electrons have been neglected completely. Bothe[3] accounts for the former effect in simple fashion by treating the electron cloud as a sphere about the nucleus of uniform charge density and with the atomic radius R as radius and $-Ze$ as total charge. Then for nuclear scattering and small values of δ, Eq. 19·11 is replaced by

$$\delta = \frac{Ze}{\Phi p}\left(1 - \frac{p^2}{R^2}\right)^{\frac{1}{2}} \quad p < R \qquad [19\cdot18]$$

For $p > R$, $\delta = 0$; the nucleus is completely shielded and does not influence the electrons passing by.

A final magnitude of importance is the effective scattering cross section of an atom for a given scattering angle δ, that is, the cross-sectional area of the beam within which the electrons are scattered through angles greater than δ. This is, by Eq. 19·17*a*, for relatively small angles and disregarding the screening effect,

$$\sigma = \sigma_n + \sigma_e = \pi p_n^2 + Z\pi p_e^2 = \frac{\pi Z^2 e^2}{\Phi^2 \delta^2}\left(1 + \frac{1}{Z}\right) \qquad [19\cdot19]$$

If account is taken of the shielding of the nucleus in accordance with Eq. 19·18, σ can obviously never exceed πR^2, since beyond R the atom does not influence the incident electrons. Accordingly, in place of increasing indefinitely with decreasing δ, the scattering cross section σ must approach the fixed value πR^2 as a limit.

The quantum-mechanical treatment of nuclear scattering leads to similar results, with the distinction that the upper limit of the scattering cross section is not a fixed characteristic of the atom, but is inversely proportional to the square of the velocity of the incident electrons. Thus Marton and Schiff[4] find from an application of Born's first-order theory of scattering to the Thomas-Fermi model of the atom that

$$\sigma_n = \frac{A}{4\pi^2}\lambda^2\left(1 + \frac{e\Phi}{mc^2}\right)^2 \quad \delta < \delta_1 \qquad [19\cdot20a]$$

$$A = 22Z^{\frac{1}{3}} \quad \text{for} \quad Z \geq 6$$

$$\sigma_n = \frac{\pi Z^2 e^2}{\Phi^2 \delta^2}\left[\frac{1 + \dfrac{e\Phi}{mc^2}}{1 + \dfrac{e\Phi}{2mc^2}}\right]^2 \quad \delta > \delta_1 \qquad [19\cdot20b]$$

[3] See W. Bothe's article in Geiger and Scheel, reference 6, pp. 1–74.

[4] See reference 3.

Here

$$\delta_1 = \frac{Z^{1/2}c}{180\left[\frac{2e\Phi}{m}\left(1+\frac{e\Phi}{2mc^2}\right)\right]^{1/2}} = \frac{2.81Z^{1/2}}{\{\Phi(1+0.978\cdot10^{-6}\Phi)\}^{1/2}} = 2.29\cdot10^{7}\cdot\lambda Z^{1/2} \qquad [19\cdot20c]$$

λ represents the de Broglie wave length of the electrons.

The cross section σ_e due to the inelastic collisions is calculated by Marton and Schiff by summing the excitation and ionization probabilities of the atom. They thus find:

$$\sigma_e = \frac{B}{4\pi^2}\lambda^2\left(1+\frac{e\Phi}{mc^2}\right)^2 \log\left[\frac{4e\Phi}{I}\left(1+\frac{e\Phi}{2mc^2}\right)\right] \qquad [19\cdot21a]$$

$$\delta < \frac{I}{\left[2e\Phi\left(1+\frac{e\Phi}{2mc^2}\right)\right]}$$

$$\sigma_e = \frac{B}{4\pi^2}\lambda^2\left(1+\frac{e\Phi}{mc^2}\right)^2 \log\left\{\frac{I}{e\Phi\left(1+\frac{e\Phi}{2mc^2}\right)\delta^2}\right\} \qquad [19\cdot21b]$$

$$\frac{I}{2e\Phi\left(1+\frac{e\Phi}{2mc^2}\right)} < \delta < \left\{\frac{I}{e\Phi\left(1+\frac{e\Phi}{2mc^2}\right)}\right\}^{1/2}$$

The constant B and I, the mean excitation energy of the atom, are functions of the atomic number. For example, for H, $B = 13.3$, $I = 13.5$ electron-volts; for C, $B = 35$, $I = 44$ electron-volts; and for Ag, $B = 200$, $I = 170$ electron-volts. In the case of metals, a further term may have to be added to σ_e to take account of the scattering by the conduction electrons.[5]

It has already been seen that electrons deflected by large angles may reasonably be regarded as having undergone only one significant scattering process, even for objects for which smaller deflections are, on the average, the result of a number of collisions giving rise to scattering angles of the same order of magnitude. It is of some importance to establish, for an object of given mass density, the least angle for which such single scattering applies. Wentzel[6] suggests that this least angle

[5] See Marton and Schiff, reference 3.

[6] See reference 7.

be just four times as large as a certain angle δ_m. The angle δ_m is such that the probability of scattering through angles larger than δ_m, as calculated by the formulas for single scattering, becomes just equal to 2. Thus, for $\delta = \delta_m$, $Nx\sigma_n = 2$ — account being taken only of nuclear scattering. By Eq. 19·19

$$\delta_m = \frac{Ze}{\Phi}\left(\frac{\pi\rho x N_A}{2A}\right)^{\frac{1}{2}} = \frac{1.4\cdot10^5}{\Phi} Z \left(\frac{\rho x}{A}\right)^{\frac{1}{2}} \qquad [19\cdot22]$$

Thus, for $\Phi = 50{,}000$ volts, $Z = 6$, $A = 12$, $\rho x = 10^{-6}$ gram per square centimeter, the angles considered must be larger than $4\delta_m = 1.9\cdot10^{-2}$ radian. It is thus evident that true single scattering takes place only in exceedingly thin specimens.

The other extreme, multiple scattering, can also be treated analytically. If each angular deflection is regarded as quite independent of the preceding ones, the total deflections produced by repeated scattering in a layer of matter must have a distribution given by Gauss' error curve:

$$I(\delta)d\Omega = (2\pi\Delta^2)^{-1} e^{-\delta^2/(2\Delta^2)} d\delta \qquad [19\cdot23]$$

Δ being the most probable total deflection experienced by the electron. The value of Δ is obtained by adding up statistically the deflections received in successive thin layers of thickness d of the specimen with a total thickness x, d being made so small that the single scattering formula for the shielded atom, Eq. 19·18, may be applied. The contribution of the scattering by the atomic electrons may, in a first approximation, be neglected. According to the theory of errors, the mean square total deflection $\overline{\delta^2}$ is equal to the sum of the mean square deflections in each layer, $\overline{\delta_d^2}$:

$$\overline{\delta_d^2} = 2\pi N d \int \delta^2 p dp \qquad [19\cdot24]$$

$$\overline{\delta^2} = \int_0^\infty \delta^2 (2\pi\Delta^2)^{-1} e^{-\delta^2/(2\Delta^2)} 2\pi\delta\cdot d\delta = 2\Delta^2 = 2\pi N x \int \delta^2 p dp \qquad [19\cdot25]$$

so that

$$\Delta^2 = \pi N x \int \delta^2 p dp \qquad [19\cdot26]$$

The integral in Eq. 19·26 must be carried out from a lower limit p_1, that is, angles which are so large that multiple scattering no longer takes place, to the upper limit R. Then

$$\Delta^2 = \pi N x \left(\frac{Ze}{\Phi}\right)^2 \left\{\log\left(\frac{R}{p_1}\right) - \frac{11}{12}\right\} \qquad [19\cdot27]$$

Empirically, the value of $\log (R/p_1) - 11/12$ is very nearly constant, independent of the velocity of the incident electrons as well as of the order number of the scattering atoms and even of the mass and charge of the incident particles (electrons or ions). Thus Bothe[7] arrives at the formula, valid also for electron velocities for which the relativistic effect cannot be neglected,

$$\Delta = \frac{8\cdot 10^5}{\Phi} \frac{\Phi + 5.11\cdot 10^5}{\Phi + 10.22\cdot 10^5} Z \left(\frac{\rho x}{A}\right)^{1/2} \qquad [19\cdot 28]$$

For more recent treatments of multiple scattering the reader is referred to the work of E. J. Williams[8] and of Goudsmit and Saunderson.[9]

The study of the region of plural scattering, of special interest in the case of electron microscopy, leads to great complications. However, it may be shown that if the fraction of electrons entering a small aperture from a portion of the object of given mass thickness is the only quantity of interest, this may be derived from the single scattering cross sections in simple manner even for object thicknesses for which multiple scattering would apply. Considering, again, the specimen layer divided into many infinitesimal layers dx, the decrease in the beam current remaining within the angular range from 0 to δ is given by

$$di = -i N dx \sigma(\delta)$$

i being the incident electron current and σ the atomic scattering cross section for angles greater than δ. Integrating this leads to a fraction of the current remaining within this angular range after passing through the specimen layer of thickness x:

$$\frac{i}{i_o} = e^{-Nx\sigma} \qquad [19\cdot 29]$$

As long as this quantity is large compared to the fraction of electrons scattered back into the angular range 0 to δ, given by Eq. 19·25 for multiple scattering:

$$\frac{i'}{i_o} = \int_0^\delta I(\delta)\cdot 2\pi\delta d\delta = 1 - e^{-\delta^2/(2\Delta^2)} \cong \frac{\delta^2}{2\Delta^2} \qquad [19\cdot 30]$$

Eq. 19·29 may be taken as a true measure of the fraction of the electrons which ultimately pass through the perture in question. This condition is satisfied for specimen mass thicknesses $\rho x < 6\cdot 10^{-5}$ gram per square centimeter (6000 A.U. for $\rho = 1$) if $\delta = 10^{-4}$ radian, for

[7] See reference 6.
[8] See reference 4.
[9] See reference 5.

$\rho x < 3{\cdot}10^{-5}$ gram per square centimeter (3000 A.U. for $\rho = 1$) if $\delta = 10^{-2}$ radian, if the accelerating voltage is 50 kilovolts and the atomic number of the specimen is $Z = 6$ (carbon). Marton and Schiff's values are used for the atomic cross sections (Eqs. 19·20 and 19·21). Since the fraction of electrons remaining within the angular range 0 to δ is of the order of 10^{-6} and 10^{-2}, being given roughly by $100\delta^2$, in the two cases cited, this range easily covers the thicknesses normally encountered in electron-microscope specimens. If the accelerating voltage is altered, conditions are only slightly changed provided that the specimen thickness is altered in direct proportion thereto.

It should be remarked that Eq. 19·29 is strictly valid only for angles δ which are so small that the scattering cross section is independent of δ. That is, by Eq. 19·21*a*, δ must be smaller than half the ratio of the mean ionization potential of the atoms to the accelerating potential of the incident electrons. For example, for $C(Z = 6)$ and $\Phi = 50{,}000$ volts, $\delta \leq 4.4{\cdot}10^{-4}$ radian. For larger angles δ, Eq. 19·29 will give too large values.

In addition to the scattering in angle, the scattering in kinetic energy of the electrons resulting from inelastic collisions affects the imaging process in the electron microscope. If the beam leaving the object and entering the objective is inhomogeneous in velocity, the objective lens, in view of its inherent chromatic aberration, cannot form a perfectly sharp image. The relation between the loss in kinetic energy and the angular deflection resulting from an inelastic collision has been given in Eq. 19·3*b*. Although this does not take account of the binding of the electrons to the atom and hence cannot claim strict validity for small angles of deflection, it may be used to determine the mean loss in kinetic energy which an electron experiences in passing through a layer of infinitesimal thickness dx:

$$d(e\Phi) = 2\pi ZN dx \int \Delta(e\Phi) p_e dp_e \qquad [19{\cdot}31]$$

Substituting from Eq. 19·14 for p_e and Eq. 19·3*b* for $\Delta(e\Phi)$ gives

$$\begin{aligned} \frac{d(e\Phi)}{dx} &= -2\pi ZNe\Phi \int_{\delta_1}^{\pi/2} \sin^2 \delta \frac{e^2 \cos \delta}{\Phi^2 \sin^3 \delta} d\delta \\ &= \frac{2\pi ZNe^8}{\Phi} \log (\sin \delta_1) \end{aligned} \qquad [19{\cdot}32]$$

The lower limit of integration δ_1 represents the minimum angle of deflection which corresponds to a possible energy transfer to an atomic electron. As Bethe[10] has shown, this minimum angle does not corre-

[10] See reference 8.

spond to a classical energy transfer equal to the excitation or ionization energy I, but to the much smaller quantity $I^2/(4e\Phi)$. That is, it is given, according to Eq. 19·3b, by

$$\sin \delta_1 = \frac{I}{2e\Phi} \qquad [19\cdot33]$$

Thus the rate of loss of energy of the electron in passing through a layer of matter is, on the average,

$$\frac{d(e\Phi)}{dx} = -\frac{2\pi ZNe^3}{\Phi} \overline{\log\left(\frac{2e\Phi}{I}\right)} \qquad [19\cdot34]$$

where the logarithmic term is suitably averaged to take account of the different ionization potentials of the atomic and conduction electrons. Integrated, this equation leads to the relation

$$\Phi_o^2 - \Phi^2 = 4\pi ZNe^2 \overline{\log\left(\frac{2e\Phi}{I}\right)} \cdot x \qquad [19\cdot35]$$

where $e\Phi_o$ is the kinetic energy of the incident electrons and $e\Phi$ the average energy of the electrons which have passed through a layer thickness x. For fast electrons the logarithmic term may be regarded as a constant, so that Eq. 19·35 takes on the simple form

$$\Phi_o^2 - \Phi^2 = \frac{a\rho Z}{A} x \qquad [19\cdot35a]$$

a is a constant which shows only a relatively slight dependence on atomic number (for Pt [78] it is found to be about four times as large as for Al [13]).[11] For a mixed substance with a relative number n_j of atoms with atomic number Z_j and atomic weight A_j, the ratio aZ/A must merely be replaced by $\sum_j (n_j a_j Z_j)/\sum_j (n_j A_j)$. Equation 19·35$a$ is known as the Thomson-Whiddington law,[11] more commonly written in the form

$$v_o^4 - v^4 = \frac{b\rho Z}{A} x \qquad [19\cdot36]$$

where v_o is the initial velocity, v the most probable final velocity of the electrons after passing through a layer of matter of thickness x. For aluminum,[12] b has a value between 5 and $10\cdot10^{42}$ in the range of 6000 to 60,000 volts, corresponding to $a = 4\cdot10^{11}$ to $8\cdot10^{11}$.

For very high voltages Eqs. 19·35 and 19·36 cease to be valid.

[11] See Whiddington, reference 9.

Bothe,[12] on the basis of a compilation of data on aluminum by Lenard,[13] finds that Eq. 19·36 is valid up to about 130,000 volts with $b = 5.4 \cdot 10^{42}$. Above 200,000 volts

$$\frac{d\Phi}{dx} = -4.56 \cdot 10^6 \text{ volts per cm} \qquad [19\cdot37]$$

yields an adequate representation of the data for aluminum. Other empirical laws[14] show a similar reduction in the dependence of the rate of energy loss on the electron velocity at high voltages.

If, in Eqs. 19·35*a* and 19·36, Φ and v are set equal to zero and the equations are solved for x, the *penetrations* of electrons of the given initial kinetic energy $e\Phi_o \cong mv_o^2/2$,

$$x_m = \frac{A\Phi_o^2}{a\rho Z} \cong \frac{Av_o^4}{b\rho Z} \qquad [19\cdot38]$$

are obtained. These are the thicknesses of matter for which the " most probable " exit velocity has just been reduced to zero. In Fig. 19·5, representing velocity distributions of very high velocity electrons (β-rays) after traversing different thicknesses of mica,[15] this most probable velocity corresponds to the abscissa of the maximum. It is seen that the velocity distribution spreads over a very large range long before the maximum penetration, as given by Eq. 19·38, has been reached.

An important practical question, arising in connection with the electron microscope, is whether the electrons which remain in a narrow angular range between 0 and δ_2 after passing through a layer of matter experience velocity losses comparable with the average velocity loss of all the electrons. It can readily be seen that the average velocity loss of the selected electrons must be smaller. Thus the effect of extending the consideration to electrons leaving in the range 0 to δ_2 is to replace the logarithmic factor in Eq. 19·34 by $\log (2e\Phi \sin \delta_2/I)$. Suppose that $\sin \delta_2$ is $3 \cdot 10^{-3}$. Although the proper average value of I, the ionization or excitation energy of the atom, to be employed in this formula is difficult to determine, it is clear that atomic electrons with an excitation energy in excess of $2e\Phi \cdot 3 \cdot 10^{-3}$, or 300 volts for $\Phi = 50{,}000$ volts, cannot contribute to the energy loss. Thus, for example, in aluminum the K-electrons play a negligible role. With I set equal to a straight average of the ionization energies of the remaining electrons (56 electron-

[12] See W. Bothe in Geiger and Scheel, reference 6, pp. 1–74.
[13] See reference 10.
[14] See Williams, reference 11.
[15] See White and Millington, reference 12.

volts), the reduction in the average energy loss of the selected electrons as compared with that for all the electrons is given by the factor

$$\frac{11 \log \left(\frac{300}{56}\right)}{2 \log \left(\frac{100{,}000}{1550}\right) + 11 \log \left(\frac{100{,}000}{56}\right)} \cong 0.2$$

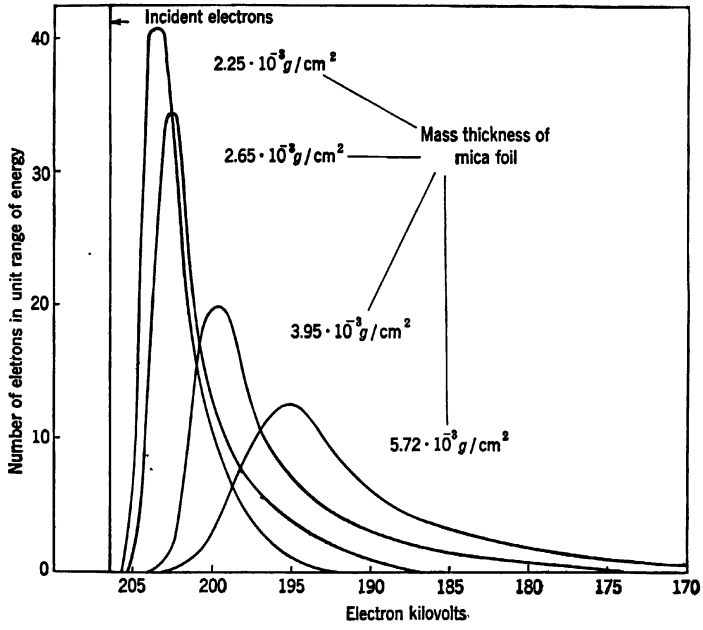


FIG. 19·5. Velocity Distributions of High-Velocity Electrons after Traversing Different Thicknesses of Mica. (White and Millington, reference 12. Courtesy of *Proc. Roy. Soc.*, London.)

Furthermore, if, as is permissible with the specimen thicknesses normally employed in the electron microscope, the electrons scattered back into the central cone can be neglected in comparison with those that have remained therein throughout, that is, with the fraction $e^{-\sigma N x}$ of the incident electrons, none of the transmitted electrons will have suffered losses in velocity if $\delta \leq \delta_1 = I/(2e\Phi)$, provided that the minimum excitation or ionization potential is substituted for I. Thus if $I = 5$ electron-volts and $\Phi = 50{,}000$ volts, the transmitted electrons will have experienced no loss in velocity if $\delta \leq 5 \cdot 10^{-5}$ radian. For larger angles δ the velocity

distribution of the electrons within the central cone will still have a " peak " corresponding to zero velocity loss.

The nearest approach to an experimental test of this question was carried out by Boochs,[16] who measured the velocity losses after passage through an aluminum foil 0.8 micron thick ($\rho x = 2 \cdot 10^{-4}$ gram per square centimeter) both for the direct beam and for electrons scattered through 2.5 degrees, the incident beam having an angular aperture of $3 \cdot 10^{-3}$ radian. He found a velocity loss of 0.8 per cent for the direct beam, of 1.0 per cent for the electrons leaving with an inclination of 2.5 degrees. This is consistent with the preceding deductions, since Boochs' foil was much too thick to permit a neglect of the electrons scattered back into the central cone. The estimates given for the reduction in velocity loss for the electrons leaving at small inclinations to the incident beam are based on the assumption that these diffusely scattered electrons can be neglected.

The nature of the velocity distribution curves in Fig. 19·5 makes it clear that even for foils considerably less than the penetration distance in thickness a large fraction of the incident electrons will not reach the back surface of the foil. There exists thus an *absorption* of electrons as well as a reduction in velocity. For this absorption Lenard[17] established the validity, within limits, of the usual absorption law

$$n = n_0 \cdot e^{-\alpha x} \qquad [19\cdot39]$$

α being proportional to the inverse fourth power of the electron velocity and approximately directly proportional to the density of the material. By making use of Terrill's measurements on aluminum,[18] covering the range between 18,000 and 40,000 volts,

$$\frac{\alpha}{\rho} = \frac{8 \cdot 10^{11}}{\Phi^2} \text{ cm}^2/\text{g} \qquad [19\cdot40]$$

Although this relation is in no sense exact (for some halogen compounds and the halogens themselves the numerical factor may exceed that given by a factor of 2 or more) it suffices to demonstrate that electron absorption plays a negligible role in determining the portion of electrons from a given object element which contributes to the formation of the image in the electron microscope. Scattering alone considered, this proportion is

$$e^{-\sigma N x} = e^{-c_a \rho x} \qquad c_a = \frac{\sigma N_A}{A} \qquad [19\cdot41]$$

[16] See reference 13.
[17] See reference 10, p. 73.
[18] See Terrill, reference 14.

On the other hand, absorption alone considered, the proportion of electrons remaining in the cone of angular aperture δ which contributes to the imaging is

$$e^{-ax} = e^{-c_a \rho x} \qquad c_a = \frac{8 \cdot 10^{11}}{\Phi^2} \qquad [19\cdot42]$$

If, for example, the specimen is carbon and $\Phi = 50{,}000$ volts, $c_s = 1.5 \cdot 10^5$ square centimeters per gram for $\delta = 3 \cdot 10^{-3}$ radian and $c_s = 4.7 \cdot 10^3$ square centimeters per gram for the very large angle $\delta = 10^{-1}$ radian (5.7 degrees). In either case $c_a = 320$ square centimeters per gram, that is, less by a factor of at least 15. It is thus seen that the effect of absorption in the specimen on reducing the intensity of the transmitted effective electron beam is quite negligible compared to that of scattering.

19·3. Image Formation with Small Aperture. The correspondence between the image formation in the electron microscope and that in the light microscope becomes closest if a narrow physical aperture is used. In treating this case it will be assumed that the full free aperture of the objective is uniformly illuminated by the electrons leaving any imaged point of the object. This may take place either because the condenser aperture, that is, the angular aperture of the beam illuminating the object, exceeds the angular aperture of the beam accepted by the objective (Fig. 19·6) or because the scattering in the object is such that the electron distribution in the narrow central cone entering the objective is uniform. In either case, for the specimen thicknesses normally encountered in electron microscopy, the relative intensity of the radiation entering the objective from different points of the object is given by

$$\frac{i}{i_o} = e^{-\sigma N x}$$

where σ is the atomic scattering cross section, N the number of atoms in unit volume of the specimen, and x the thickness of the specimen at the point in question. This equation simultaneously determines the relative concentration of electrons in different parts of the image and hence the contrast of the image.

Since, as Scherzer has pointed out,[19] an electron wave after passing through a single layer of atoms shows random phase variations averaging $\frac{1}{100}$ cycle, the electron radiation emitted by different points of a specimen mounted on a supporting film may quite generally be regarded as incoherent, so that an object in the electron microscope normally acts as though it were self-luminous.

[19] See reference 15.

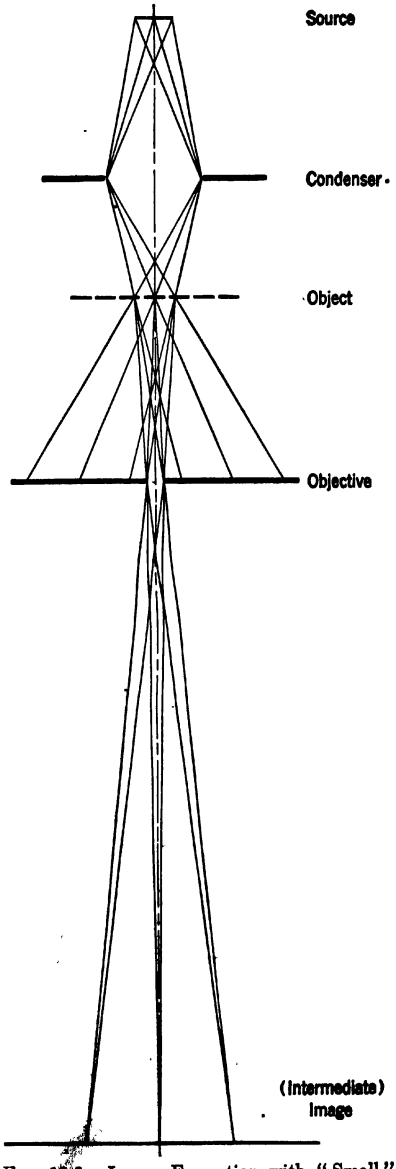


FIG. 19·6. Image Formation with "Small" Aperture.

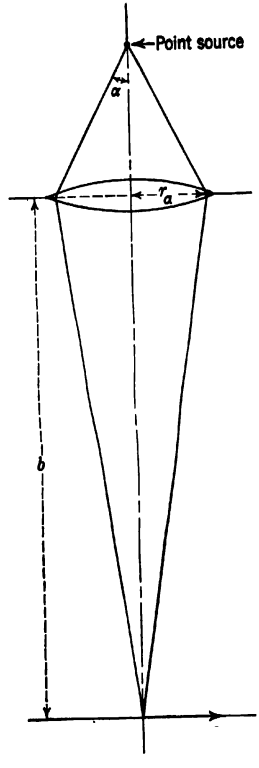


FIG. 19·7. Notation Employed in the Derivation of the Diffraction by the Objective Aperture.

Consider (Fig. 19·7) a point source of radiation in the object plane on the axis of the objective. The divergent radiation from this point will be rendered nearly parallel in passing through the lens and, at the same time, will be restricted by the aperture of the objective. To a certain extent the radiation will be diffracted at the edge of the aperture. It is convenient to regard this process as taking place in image space, at the exit pupil (that is, the image of the actual aperture), whose radius will be denoted by r_a. Then the theory of Fraunhofer diffraction may be applied to the nearly parallel ray pencil.[20] For aberration-free imaging this leads to an intensity distribution in the plane of the image of the point source

$$I = \left(\frac{\pi r_a^2}{\lambda b f}\right)^2 \left\{ \frac{2J_1\left(\frac{2\pi\rho}{\lambda}\frac{r_a}{b}\right)}{\frac{2\pi\rho}{\lambda}\frac{r_a}{b}} \right\}^2 \qquad [19\cdot43]$$

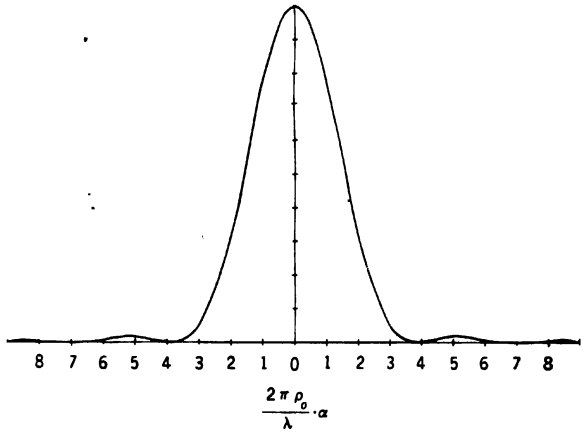


FIG. 19·8. Diffraction Pattern Formed as Image of Point Source.

Here ρ is the distance from the Gaussian image point, b the separation of the image from the exit pupil, f the focal length of the lens, and λ the wave length of the radiation. J_1 is the Bessel function of the first order. Figure 19·8 shows a plot of the intensity distribution in Eq. 19·43. It is seen to become zero for

$$\frac{2\pi\rho}{\lambda}\frac{r_a}{b} = 1.220\pi, 2.233\pi, 3.238\pi, 4.250\pi, \cdots \qquad [19\cdot43a]$$

[20] See Born, reference 16, Chapter 4, section 49.

For a sufficiently small objective aperture to make the effect of the lens aberrations negligible, every point in the object gives rise to such a diffraction figure or *Airy disk* in the image plane. Consider two object points a small distance d apart. The centers of their diffraction disks

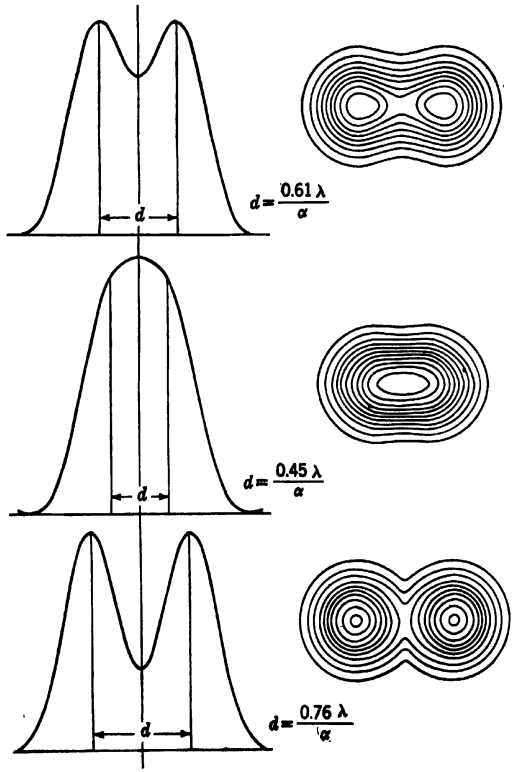


FIG. 19·9. Intensity Distribution in Aberration-Free Image of Two Neighboring Point Sources.

will be a distance Md apart, M being the magnification of the image. If Md is greater than the distance between the center and the first minimum (or zero) of the diffraction figure, a dip will appear between the two intensity peaks (Fig. 19·9), facilitating their separation. However, if the two disks are closer together, the dip in intensity eventually disappears and the two diffraction figures appear to coalesce, a maximum of intensity existing halfway between the centers of the two diffraction patterns. Rather arbitrarily, the condition that the peak of one diffrac-

tion figure coincides with the first minimum of the other is used to define the limit of resolution, that is, to establish the least separation of two points of the object which will enable the observer to recognize the two points as two separate entities. The actual minimum distance which can be separated depends, of course, on the practice of the observer and his previous knowledge with regard to the nature of the specimen.

If α is the half angle subtended by the aperture at the object — in short the *aperture* of the objective — and n_o is the index of refraction of the medium in which the object is placed, n_i being that of image space, the theorem of Helmholtz-Lagrange demands that

$$n_o \sin \alpha = \frac{M n_i r_a}{b}$$

so that the condition that the diffraction maximum for one point coincide with the first minimum for the other takes the form

$$\frac{2\pi M d}{\lambda_i} \frac{n_o \sin \alpha}{n_i M} = 1.220\pi$$

or

$$d = \frac{0.61\lambda_o}{\sin \alpha} \qquad [19\cdot44]$$

Here $\lambda_o = (n_i/n_o)\lambda_i$ is the wave length of the imaging radiation at the object.

The order of magnitude of the limitation of the resolving power by diffraction effects can also be deduced from an application of the quantum-mechanical uncertainty relation.[21] According to this relation, in a rather specialized form, the component of momentum of a particle in the x-direction can be determined only within an accuracy $\Delta(mv_x)$ which is related to the accuracy Δx with which its coordinate x can be measured by the relation

$$\Delta(mv_x)\cdot\Delta x \cong h$$

where h is Planck's constant. Consider now an electron of energy $mv^2/2 = e\Phi$ leaving a point of the object (Fig. 19·10). It will pass through the lens and impinge on the viewing screen near the Gaussian image point, provided that its lateral component of momentum lies in the range from 0 to $\pm mv \sin \alpha$. Thus, if its impact is observed on the screen, there is an uncertainty in its component of momentum at the time of leaving the object equal to $2mv \sin \alpha$. Accordingly, its place of arrival must show an uncertainty in the x-direction such that from it

[21] See section 10·2.

its origin in the object can be determined only within an accuracy

$$\Delta x \cong \frac{h}{\Delta(mv_x)} = \frac{h}{2mv \sin \alpha} \doteq \frac{\lambda}{2 \sin \alpha} \qquad [19\cdot44a]$$

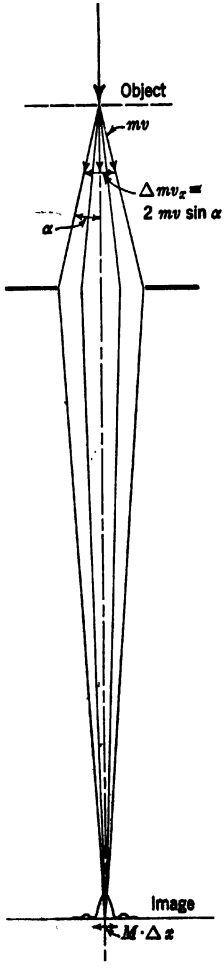


FIG. 19·10. Derivation of Limit of Resolving Power from Uncertainty Relation.

This equation is substantially identical with Eq. 19·44.

If the aperture of the objective is sufficiently small that

$$\frac{0.61\lambda}{\alpha} \gg Cf\alpha^3$$

f being the focal length of the objective and C the coefficient of spherical aberration, that is, if the radius of the circle of confusion due to the aperture defect is much smaller than the least resolvable distance as determined by diffraction only, the resolution is given simply by Eq. 19·44, just as for a perfectly corrected light objective. However, even if satisfactory apertures of arbitrarily small size could be made and would be free from other drawbacks, it would yet be desirable to make the aperture large enough for the optimum resolution of which the objective is capable to be realized. Accordingly, the angular aperture should be increased until the improvement brought about by the reduction in size of the diffraction pattern is overbalanced by the deterioration of the image from the increase in the circle of confusion due to the aperture defect.

A thorough analysis of this problem requires the calculation of the diffraction pattern obtained for waves which, instead of converging on the Gaussian image point, converge on a caustic characteristic of the spherical aberration of the lens in question. Calculations for such aspheric waves have been carried through by Picht.[22] Figure 19·11 shows the diffraction pattern for a point object formed by a lens with the indicated amount of spherical aberration, calculated with the aid of Picht's formulas. Al-

[22] See reference 17, sections 7 and 14.

though the general character of the distribution has not been changed by the presence of the spherical aberration, a larger fraction of the radiation is spread over the outer portions of the pattern and the minima are slightly less pronounced, decreasing not all the way to zero.

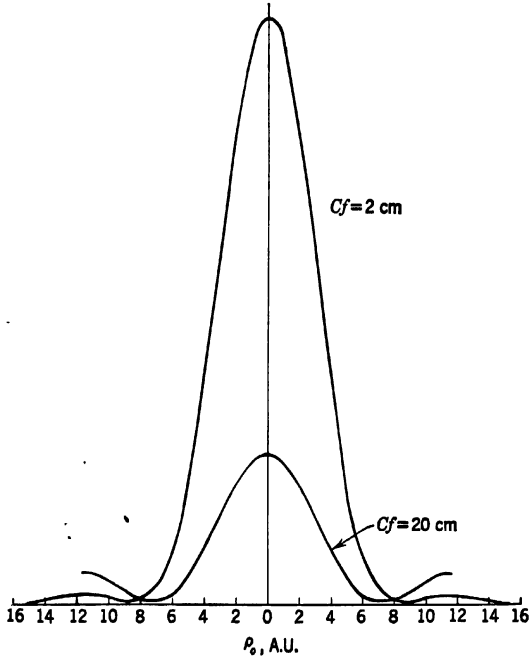


FIG. 19·11. Image of Point Source, Taking Account of Diffraction and Spherical Aberration ($\lambda = 5.35 \cdot 10^{-10}$ cm).

Even without proceeding to an actual evaluation, the variation of the resolving power, as determined by diffraction and spherical aberration, with the objective aperture α, the wave length λ (and hence the accelerating voltage), and the spherical aberration coefficient Cf can be deduced from the general functional character of Picht's formulas. The intensity distribution in the diffraction pattern may be written

$$I = \left(\frac{1}{\lambda Cf}\right) F\left(\frac{Cf}{\lambda}\sin^4\alpha; \frac{\rho_o}{\lambda}\sin\alpha\right) \qquad [19\cdot45]$$

ρ_o is the separation of the point of reference from the Gaussian image point, referred to the object. It is seen that, as long as the quantity

$(Cf/\lambda)\sin^4\alpha$ is left unchanged, the only effect of changing Cf, λ, or α is to cause a contraction or expansion of the pattern about the Gaussian image point with a simultaneous variation in the absolute values of the intensities. Accordingly, if the intensity distributions are known for all values of α for specific values of Cf and λ, the distributions for arbitrary α, Cf, and λ can be deduced from these in simple manner.

Suppose that, for some fixed values of λ and Cf, the intensity distributions had been determined and that from them, with the aid of some arbitrary criterion, the least distances $d(\sin\alpha)$ between two object points recognized in the image as separate entities had been derived. For a particular value α_m, $d(\sin\alpha)$ will attain a minimum value $d(\sin\alpha_m) = d_{\min}$. This aperture α_m would be the most favorable aperture for the given values of the wave length and spherical-aberration coefficient.

Assume, next, that the wave length λ is increased by a factor K, and Cf is increased by the factor k. Then the variation in intensity as given by Eq. 19·45 will be unchanged if $\sin\alpha$ is increased by a factor $(K/k)^{1/4}$, with the exception of a simultaneous expansion outwards by a factor $K\cdot(K/k)^{-1/4} = K^{3/4}k^{1/4}$ and a reduction in the absolute values of the intensities by the factor $1/(Kk)$. The resolution as a function of $\sin\alpha$ will thus now be given by

$$d'(\sin\alpha) = K^{3/4}k^{1/4}d\left(\left[\frac{K}{k}\right]^{-1/4}\sin\alpha\right)$$

This will be a minimum for

$$\sin\alpha = \left(\frac{K}{k}\right)^{1/4}\sin\alpha_m$$

attaining a value

$$d'_{\min} = K^{3/4}k^{1/4}d_{\min}$$

It follows from these two equations that, quite generally, the optimum aperture and the least resolvable distance are given by

$$\sin\alpha_{\text{opt}} = c'\left(\frac{\lambda}{Cf}\right)^{1/4} \qquad [19\cdot46]$$

$$d_{\min} = c''\lambda^{3/4}(Cf)^{1/4} \qquad [19\cdot47]$$

The constants c' and c'' are general constants which could be determined by a series of calculations of the intensity distributions for fixed Cf and λ, as indicated above.

In the absence of these calculations, attempts at estimating the resolution have been made on the basis of simple assumptions. Thus Rebsch[23]

[23] See reference 18

places the least resolvable distance equal to the sum of the limit of resolution for diffraction only and the limit of resolution when the aperture defect alone is considered. The latter, again, requires definition. If the aperture defect alone is present there is always a peak of infinite intensity at the Gaussian image point of a luminous point object, the distribution being given by

$$J(\rho_o) = \frac{I}{3(Cf)^{\frac{3}{4}}\rho_o^{\frac{1}{2}}} \qquad [19\cdot48]$$

where I is the emission per unit solid angle of the object point. Similarly, for an object of finite size the variation in intensity at the Gaussian image of the edge of the object has infinite slope.[23] Hence there will always be a minimum of intensity between the Gaussian images of two objects. However, an increasing proportion of the light from a given luminous object will spread out into its surroundings, decreasing the contrast and ease of recognition. Hence, somewhat arbitrarily, Rebsch sets the limit of resolution due to the aperture defect equal to the width d_s of a luminous strip in the object, for which just 50 per cent of the radiation accepted by the objective falls on the Gaussian image of the strip. This is equal to $0.4Cf\alpha^3$. The condition that

$$d = 0.4Cf\alpha^3 + \frac{0.6\lambda}{\alpha} \qquad [19\cdot49]$$

be a minimum, that is,

$$1.2Cf\alpha^2 - \frac{0.6\lambda}{\alpha^2} = 0$$

leads to

$$\alpha_{\text{opt}} = \left(\frac{\lambda}{2Cf}\right)^{\frac{1}{4}} \qquad c' = 0.84 \qquad [19\cdot50a]$$

$$d_{\min} = 0.95\lambda^{\frac{3}{4}}(Cf)^{\frac{1}{4}} \qquad c'' = 0.95 \qquad [19\cdot50b]$$

The coefficients in these equations are only slightly dependent on the criterion selected for determining the effect of the spherical aberration on the resolution. According to Rebsch, the lowest obtainable value of C may be estimated at 0.25.[24] For an objective with the very short focal length of 1 millimeter the best attainable values of α_{opt} and $d_{\min}$ would thus be according to Eqs. 19·50, it being assumed that $\lambda = 5.35\cdot10^{-10}$ centimeter (50,000 volts),

$$\alpha_{\text{opt}} = 10^{-2} \text{ radian} \qquad d_{\min} = 4\cdot10^{-8} \text{ cm} = 4 \text{ A.U.}$$

[24] See section 17·2.

Available data on the aperture defects and dimensions of the best magnetic microscope objectives in use[24] at present suggest that for these Cf may have values lying between 0.2 and 2 centimeters, signifying values of α_{opt} between $6 \cdot 10^{-3}$ and $3 \cdot 10^{-3}$ radian and of d_{min} between 7 and 13 A.U.

The projector acts purely as an enlarging lens without influence on the image quality or resolution near the center of the image. Its aperture is invariably very large compared to that of the beams leaving the objective and its spherical aberration, for the fraction of the aperture utilized, is negligible. The effect of the field aberrations, in particular the distortion, of the projector has been considered in sections 17·4 and 17·5. With relatively large image fields the distortion of the projector may cause a slight falling off of the intensity near the edge of the image.

With a small physical aperture in the objective, pictures with high contrasts are obtained even with very small object thicknesses. Thus, with $\alpha = 3 \cdot 10^{-3}$, $\Phi = 50{,}000$ volts, and an object of density unity and with an effective atomic number corresponding to that of carbon, the beam current is reduced by a factor

$$e^{-1.5 \cdot 10^5 x}$$

in passing through a thickness x. Thus a difference in thickness in the object equal to $5 \cdot 10^{-6}$ centimeter (500 A.U.) will correspond to a factor of 2 in the brightness of the fluorescent screen or the number of electrons arriving at the photographic plate, a difference in thickness of 1500 A.U. to a factor of 10. With a narrow aperture there is a strict correspondence between object mass thickness and illumination of the image screen or plate in accordance with Eq. 19·29 (assuming that crystalline reflections [section 19·9] are absent) except in a region about the boundaries of object portions of constant mass thickness of a width of the same order as the resolving power of the instrument.

A difficulty of rather fundamental character arises in the employment of very small apertures from the image field exerted by the aperture edge on a passing electron. This image field is given simply by $e/(4r^2) = 3.6 \cdot 10^{-8}/r^2$ volts per centimeter, where r denotes the distance between the electron in question and the wall of the aperture bore facing it. If the lateral component of velocity acquired as a result of the action of the image force is $(2e\Phi_r/m)^{1/2}$ and the length of the channel is l, the equality between the impulse and acquired momentum ($ft = mv$) leads to

$$\frac{e^2 l}{4r^2\left(\frac{2e\Phi}{m}\right)^{1/2}} = (2em\Phi_r)^{1/2} \qquad \left(\frac{\Phi_r}{\Phi}\right)^{1/4} = \frac{el}{8r^2\Phi}$$

This results in an apparent displacement d of the object point:

$$d = f\left(\frac{\Phi_r}{\Phi}\right)^{1/2} = \frac{elf}{8r^2\Phi} = 1.8\cdot10^{-8}\frac{fl}{r^2\Phi}$$

For example, for $f = 0.3$ centimeter, $l = 10^{-2}$ centimeter, $r = 10^{-4}$ centimeter, and $\Phi = 50{,}000$ volts,

$$d = 10^{-7} \text{ cm (10 A.U.)}$$

All electrons passing within a micron or less from the wall of the bore (assuming that its diameter is materially larger than 2 microns) will thus appear to arise from points at distances of over 10 A.U. from the true object point. For a 20-micron aperture of the given thickness, 20 per cent of the electrons are deflected to this extent. If the focal length f of the objective is reduced to 0.1 centimeter and the aperture diameter is reduced in proportion, all other quantities remaining unchanged, a third of the electrons experience deflections giving rise to apparent deviations from the object point in excess of 10 A.U. Thus, if very narrow apertures are employed, it is advisable to make the aperture diaphragm as thin as is consistent with adequate opacity to electrons and mechanical strength.

19·4. Image Formation with Large Aperture. Although the employment of narrow objective apertures is advantageous in a number of respects, most electron microscopes in use at the present time have objective apertures much too large to act as true limiting apertures. Frequently, in fact, the pole pieces themselves furnish the only physical aperture of the objective.

The reasons for this development are primarily practical. First of all, if an insulating film or insulating particles adhere to the edges of a narrow aperture, they will become charged under the impact of the beam and seriously disturb the image. The absence of a narrow aperture does away with the necessity of removing and cleaning it frequently and makes the operation of the instrument somewhat less critical. Furthermore, with such an aperture focusing is difficult, on the one hand because of the impossibility of greatly increasing the brightness of the image for visual focusing by changing the condenser setting, and on the other hand, because of the very great depth of focus regardless of the condenser setting. Finally, it is important to choose a size for a narrow aperture which will correspond to α_{opt}, a quantity frequently not known in advance. With a large, or no, aperture in the objective and a practically parallel illuminating beam, the effective aperture, determining the resolution, will, for reasonably thin objects, automatically assume a value close to the optimum aperture, as will be shown below.

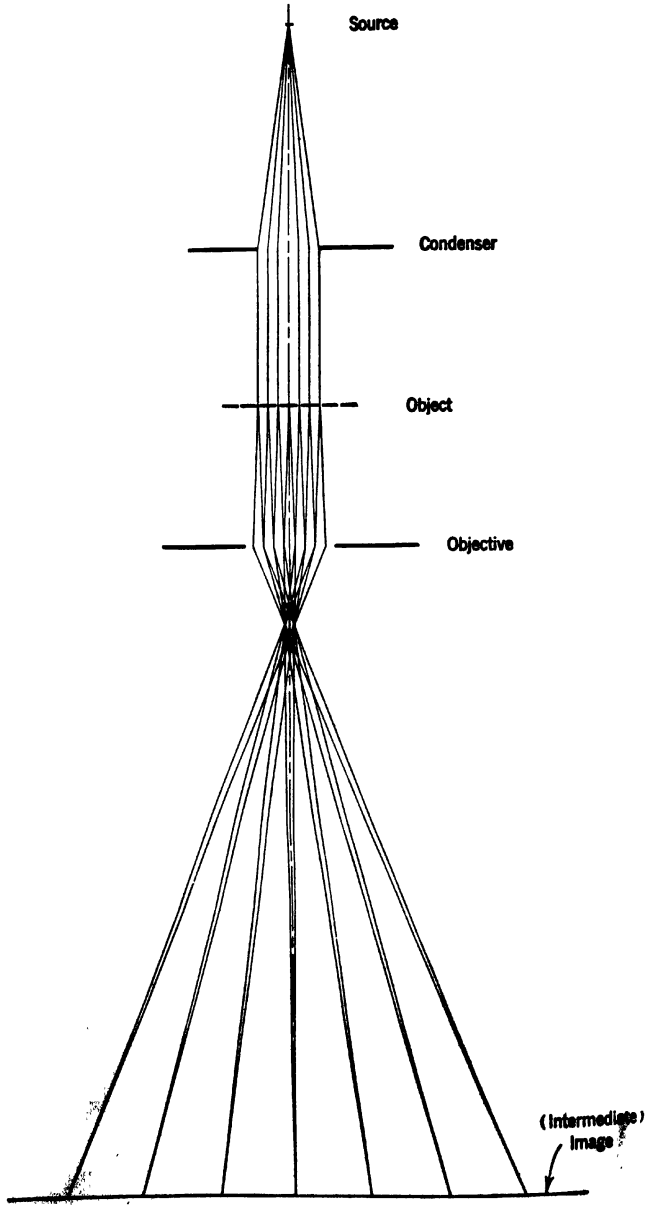


FIG. 19-12. Image Formation with a Large Aperture.

Figure 19·12 shows the imaging process with a large aperture. It is seen that the effective angular aperture of the imaging beams is determined not by the size of this aperture, but by the degree of scattering in the object. It is a familiar fact that if coherent radiation of any kind strikes a periodic structure, such as a line grating with line separation d, part of the incident radiation is scattered, giving rise to *diffracted* beams. Abbe's theory of the imaging of nonself-luminous objects by the microscope[25] teaches that at least one of the two first diffraction beams in addition to the primary beam must pass through the objective in order to give rise to an image of the structure. This image is the result of the interference of the direct beam with the diffracted beams in the image plane and will be more faithful in proportion as a larger number of diffracted beams participate in its synthesis.

The angles δ which the diffracted beams form with the primary beam are given by

$$\sin \delta = \frac{n\lambda}{d} \qquad [19·51]$$

so that the condition that at least one diffracted beam be admitted in addition to the central beam takes the form, for illumination parallel to the axis,

$$d \geq \frac{\lambda}{\sin \alpha} \qquad [19·52]$$

and, for oblique illumination,

$$d \geq \frac{\lambda}{2 \sin \alpha} \qquad [19·53]$$

These relations are generally used to define the resolving power of a microscope for nonself-luminous objects. The result is seen to be essentially the same as for the imaging of self-luminous point objects (Eq. 19·44).

If the structure of the object is aperiodic and the illuminating beam is incoherent, scattering over a range of angle of the order λ/d, where d is a separation of two elements of the structure, still takes place, even though no well-defined diffraction beams are formed. The undeflected beam, together with the radiation scattered into a cone of half angle α_{opt}, Eq. 19·46, is mainly distributed over a circular region of radius $M \cdot d_{min}$, Eq. 19·47. The remaining scattered radiation is spread, approximately in accordance with Eq. 19·48, over a range rapidly increasing with

[25] See Abbe, reference 19.

scattering angle:

$$J(\rho_o) = \frac{I\left(\left[\frac{\rho_o}{Cf}\right]^{1/3}\right)}{3(Cf)^{2/3}\rho_o^{4/3}} \qquad [19\cdot54]$$

$I(\delta)$ being the intensity scattered by the object into unit solid angle. If $I(\delta)$ should obey the Rutherford scattering law, that is, be proportional to the inverse fourth power of δ,

$$J(\rho_o) = \frac{\text{const } (Cf)^{4/3}}{\rho_o^{8/3}}$$

which represents a very rapid decrease of intensity. The electrons scattered through angles larger than α_{opt} therefore do not, normally, appreciably affect the resolution of the image, but merely spread an " electron fog " of rapidly decreasing intensity over the neighborhood of the " image disk " of radius $M \cdot d_{min}$.

The presence of spherical aberration thus has the effect of producing image contrasts even if no electrons are held back by an aperture and none are absorbed in the object. Figure 19·13 shows schematically (each line representing an equal number of electrons) how the denser portions of the object, scattering electrons on the average through larger angles, contribute, owing to spherical aberration, to the intensity in the image of the surrounding lighter portions. It is also evident from the figure that with this type of imaging there can be no exact correspondence between mass density of the object and intensity of the image. Thus the central portion of a large uniformly dense region of the object will be imaged as brightly as the central portion of a large less dense region. Strong contrasts appear only near the edges.

The condition can best be visualized by constructing the intensity variation in the image which must be expected for a number of simple examples. This has been done in Fig. 19·14*a*, *b*, and *c*. In each case the formulas of Mott and Massey for the elastic and inelastic scattering of electrons by hydrogen[26] have been applied for determining the scattering by the specimen. At the image plane the electrons from any point are assumed to be distributed according to Eq. 19·54 for $\rho_o > d_{min}$ (Eq. 19·50*b*), ρ_o being the distance of the reference point from the Gaussian image point divided by the magnification. For $\rho_o < d_{min}$ the function

$$J(\rho_o) = \frac{AJ_1^2\left(\frac{3.832\rho_o}{d_{min}}\right)}{\rho_o^2} + \frac{I\left(\left[\frac{0.98\lambda}{Cf}\right]^{1/4}\right)}{2.8Cf\lambda} \qquad [19\cdot55]$$

[26] See reference 2, p. 174.

is taken to represent the distribution of electrons. Here the first term represents the intensity distribution in an aberration-free diffraction disk with a separation of the maximum and first minimum equal to $d_{\min}$, the resolving power when the lens has a physical aperture corresponding to α_{opt} (Eq. 19·50a and b). The second term is added so as to

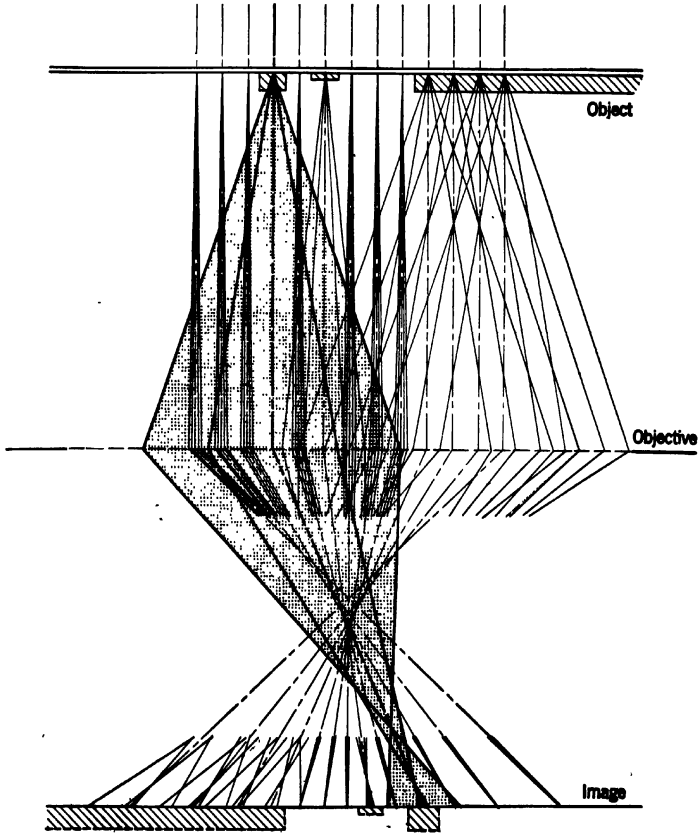


FIG. 19·13. Formation of Image Contrasts by Spherical Aberration of Objective.

make the distribution of the electrons about the Gaussian image point continuous (Fig. 19·14a). The factor A, finally, is chosen so that the total fraction of electrons falling on the disk $\rho_o < d_{\min}$ becomes equal to $e^{-N\sigma x}$, where σ is the atomic scattering cross section for $\delta_o = (d_{\min}/[Cf])^{\frac{1}{4}}$. The factor $e^{-N\sigma x}$ may be regarded as the relative contrast factor for narrow aperture operation.

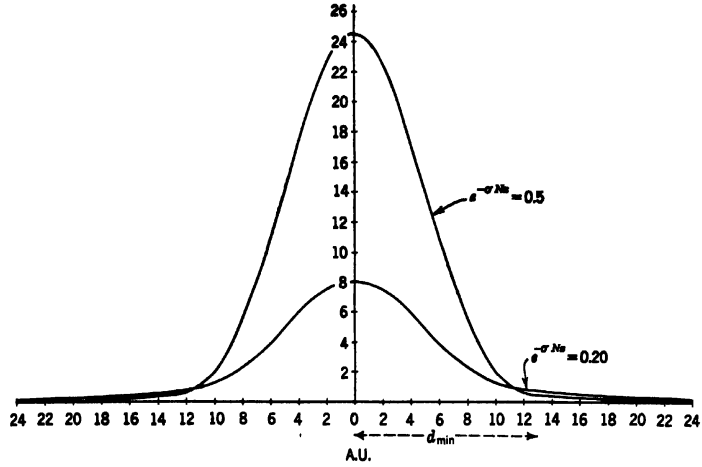


FIG. 19·14*a*.

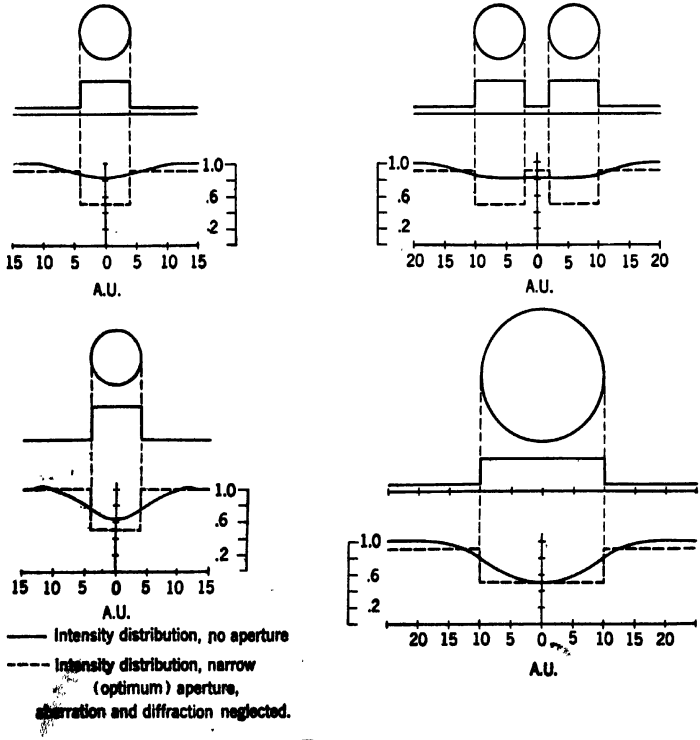


FIG. 19·14*b*.

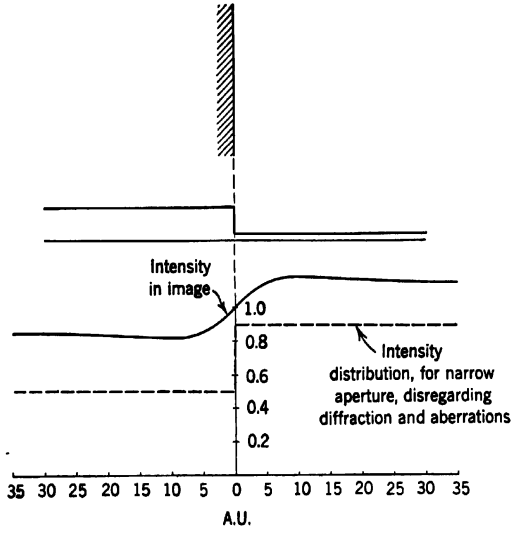


FIG. 19·14c.

FIG. 19·14. Intensity Variation in the Images of Disks and Straight Edge Formed without Limiting Aperture (Cf = 2 cm, $\lambda = 5.35 \cdot 10^{-10}$ cm).

(*a*) Assumed Intensity Distribution about the Image of a Point Source.
(*b*) Intensity Distribution in the Images of Small Disks on a Lighter Background.
(*c*) Intensity Distribution across Image of Straight Edge Separating Two Regions of Different Thickness.

In Fig. 19·14*a–c* the values $\lambda = 5.35 \cdot 10^{-10}$ (50,000 volts), $Cf = 2$ centimeters, and, hence, $d_{\min} = 13$ A.U. and $\delta_o = 4 \cdot 10^{-3}$ radian, have been assumed. The distributions have been calculated for the limiting case of no aperture restriction, which, owing to the very rapid variation of ρ_o with the scattering angle, is frequently closely approached in practice. As examples small circular disks on a lighter and on a heavier background have been chosen, the dimensions being comparable with the resolving power ($d_{\min}$), and a straight edge separating two regions of different uniform thicknesses. The disks show the decrease in contrast with size as well as the possibility of attaining, in fact, a resolution nearly equal to the optimum for a small limiting aperture. The straight-edge diagram indicates the loss in contrast arising from the absence of an aperture.

If the objects in the field (or a certain considerable fraction of them) are quite thick, the scattering function here assumed becomes inadequate. With the approach to multiple scattering, the decrease in

intensity with increasing scattering angle becomes less, and, consequently, the electrons from any portion are spread over a larger and larger area of the image. Presently, for very thick objects, an " electron fog " covers the whole field, minimizing the contrast. Thus for very thick specimens a very fine physical objective aperture is most advantageous.

19·5. Contour Effects. It is seen that in the absence of a physical aperture in the objective there is no simple correlation between image brightness and object mass thickness. Near the dividing line between two regions this is the case even if a physical aperture is employed. If the physical aperture is considerably larger than that which would corre-

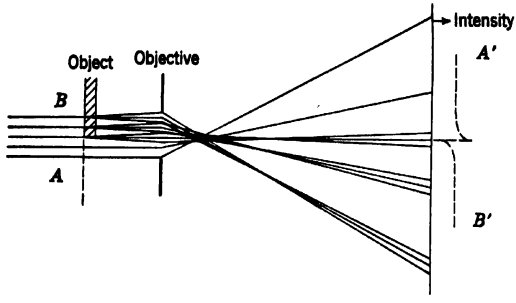


FIG. 19·15. Origin of Straight-Edge Contour Due to Spherical Aberration (Very Wide Objective Aperture).

spond to the optimum value of the angular aperture, α_{opt}, the region over which there is little correspondence between brightness and thickness extends far beyond a strip comparable in width with the resolving power of the instrument.

Consider, in particular, a region B of uniform mass thickness adjoining a wholly transparent region illuminated by a parallel beam of electrons. Leaving diffraction effects out of account, the electrons passing through the transparent region fill up uniformly the corresponding region in the image plane. Those striking the region B, on the other hand, are scattered (Fig. 19·15). Those which are scattered through angles in excess of α, the largest angle accepted by the physical aperture of the objective, do not contribute to the image. The remainder would be uniformly distributed over the region B' of the image, that is, the Gaussian image of the dense portion of the object, if it were not for the aberrations of the objective. The spherical aberration of the objective, in particular, causes a certain fraction of the scattered electrons to fall on the bright part of the image, A', up to a distance $y_{max} = C_f\alpha^3$ from the edge. If

$I(\delta)$ is the number of electrons scattered by a unit area of the dense part of the object into a unit solid angle, the distribution of scattered electrons over the clear part of the image will be given by

$$J(y) = 2\int_{\left(\frac{y}{Cf}\right)^{1/4}}^{\alpha} I(\delta) \arccos\frac{x}{Cf\delta^3}\,\delta d\delta \qquad [19\cdot56]$$

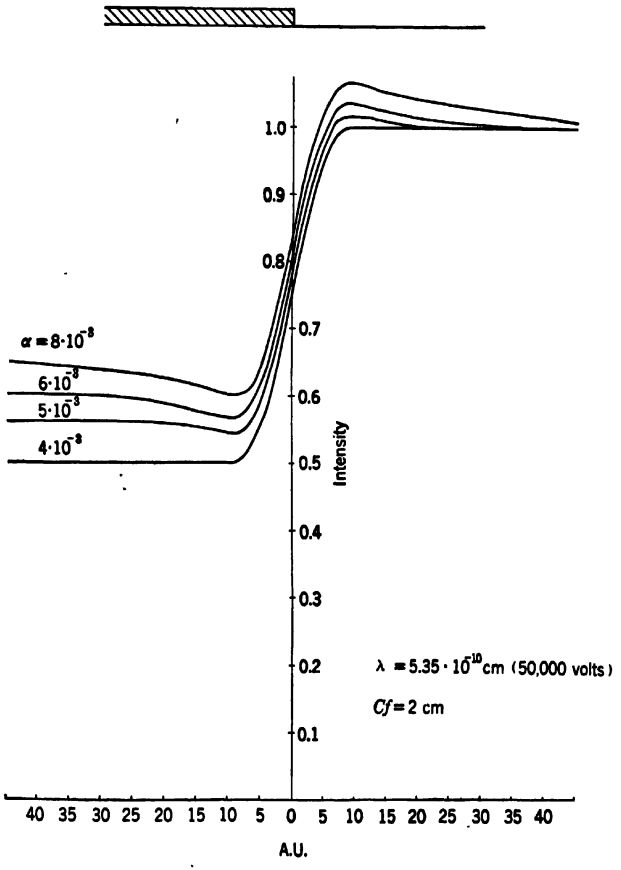


FIG. 19·16. Straight-Edge Contours Due to Spherical Aberration (and Diffraction) for Different Physical Objective Apertures.

In symmetrical manner, the same number of electrons are subtracted from the image of the dense part of the object, giving rise to a bright and a dark band on the opposite sides of the image of the edge with a conse-

quent enhancement of the contrast. This is shown for different values of α in Fig. 19·16. Other things remaining equal, an increase in α at the same time decreases the contrast and increases the dimensions of the band. In practice the sharp points at the geometrical image of the edge are rounded off, in part owing to diffraction, in part owing to the imperfect parallelism of the incident electron beam.

An enhancement of contrast, such as is shown in Fig. 19·16, takes place whenever either the lighter portion of the object is completely transparent, that is, free of matter, or if both portions are thin enough for single scattering to take place. With the approach to multiple scattering, the

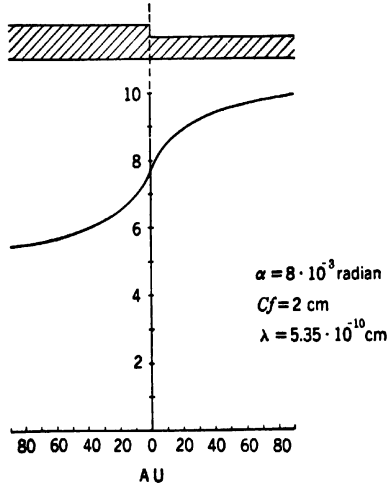


FIG. 19·17. Intensity Distribution across Border between Two Relatively Dense Portions of Specimen (Multiple Scattering.)

number of electrons scattered by the lighter portion of the object into the image of the denser portion eventually outbalances the number received by the lighter part of the image from the denser part of the object, resulting eventually in a decrease in contrast, such as is shown in Fig. 19·17.

Chromatic aberration, either due to a fluctuation of the applied voltage or to a loss in velocity of the electrons passing through the dense part of the object, has qualitatively the same effect on the appearance of an edge in the image as spherical aberration. In the absence of spherical aberration it would give rise to a contour band increasing in width in proportion to α in place of to α^3. Under normal circumstances, however, it is likely to have a smaller effect than spherical aberration

and will merely modify the intensity distribution in the contour slightly.

It is interesting to examine the behavior of these contour bands as the focus of the instrument is changed. It is seen in Fig. 19·18 that the con-

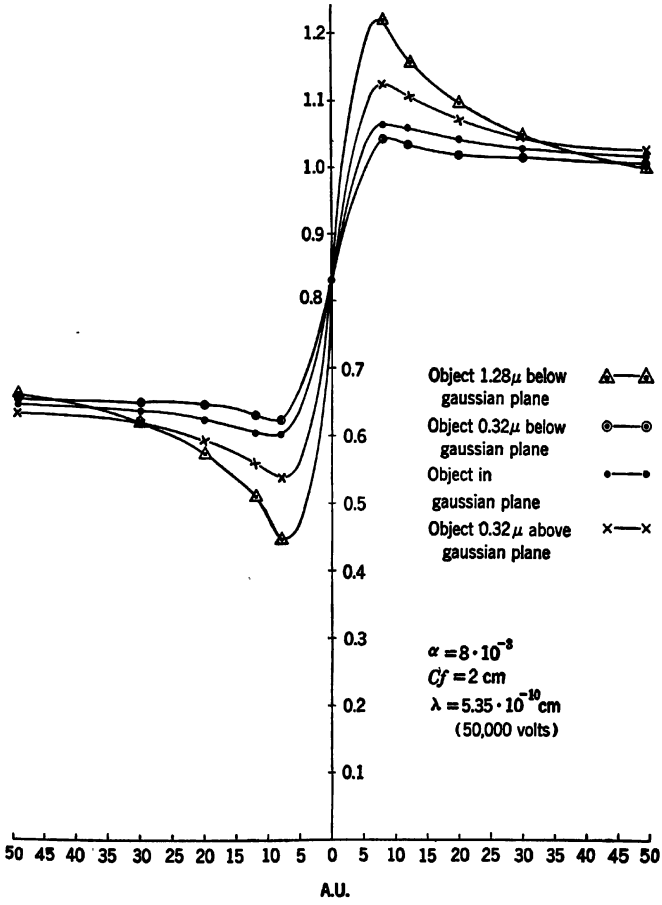


FIG. 19·18. Variation in Intensity Distribution in Contour with Focusing.

trast becomes a minimum for points intermediate between the paraxial and marginal focus, that is, for best focus. This fact is of value in determining the condition of best focus when working with a wide objective aperture. The figure again applies to the case of a completely transparent region adjoining a denser region.

For large degrees of defocusing the influence of the lens aberrations and diffraction at the objective aperture becomes secondary. The bright and dark bands on either side of the geometric image of the edge are determined primarily by the actual distribution of electrons below or by the virtual distribution above the specimen in the plane which is

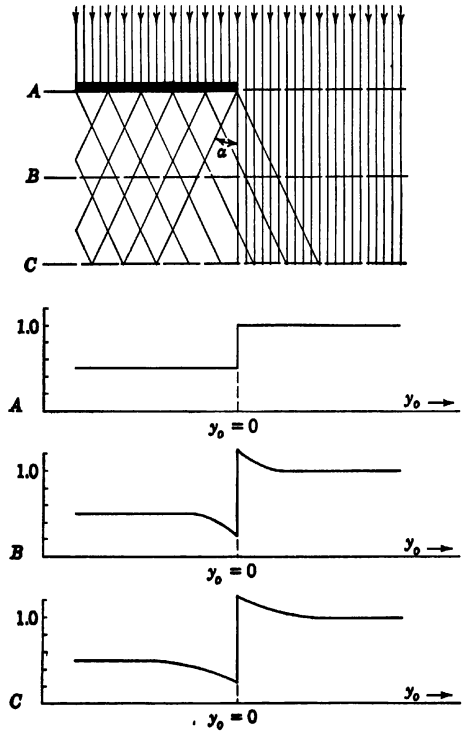


FIG. 19·19. Electron Distribution below Edge of Specimen Illuminated with a Parallel Beam of Electrons.

conjugate to the plane of the recorded image. Only those electrons which are transmitted with a path inclination less than the objective aperture α are considered. In Fig. 19·19 a parallel beam of electrons is incident on the edge of a specimen suspended in free space. It is assumed that half the incident electrons leave the specimen with inclinations less than α, that is, are accepted by the objective. Consider a plane at distance l below the edge. Inspection of the figure reveals that, taking account only of electrons which ultimately reach the objective

and contribute to the image, electrons scattered by the specimen supplement the electron concentration in a band of width

$$y_o = l\alpha \tag{19·57}$$

beyond the shadow of the specimen edge. These electrons are, in effect, contributed from a band of equal width within the shadow of the specimen.

The distribution of electrons within the band is readily determined if the simple assumption is made that the electrons transmitted by the specimen fill out uniformly the cone accepted by the objective. If the transmission of the specimen into the cone of vertex angle α is k_2, that of the free region being $k_1 = 1$, the distribution of the electrons a distance l below the edge is

$$\begin{aligned} J(y_o) &= \frac{2k_2}{\pi\alpha^2}\int_{y_o/l}^{\alpha} \arccos\frac{y_o}{l\delta}\,\delta d\delta + 1 \\ &= k_2 f\left(\frac{y_o}{l\alpha}\right) + 1 \qquad 0 < y_o < l\alpha \end{aligned} \tag{19·58}$$

$$J(y_o) = -k_2 f\left(\frac{y_o}{l\alpha}\right) + k_2 \qquad -l\alpha < y_o < 0 \tag{19·59}$$

with

$$f(x) = \frac{1}{\pi}[\arccos x - x\,(1 - x^2)^{1/2}] \tag{19·60}$$

This distribution, in two planes below the specimen, is represented in Fig. 19·19. If the electron distribution in the cone accepted by the objective is nonuniform, that is, drops off from the center toward the edge, the sole change in the distribution curves will be a more rapid falling off of the concentration from its maximum value $1 + k_2/2$ to 1 at the edge of the band on the bright side and a corresponding more rapid increase of the concentration from the minimum $k_2/2$ to k_2 on the dark side. In any case, the contrast at the edge is enhanced from a ratio $1 : k_2$ to $(2 + k_2) : k_2$, regardless of the distance l between the focusing plane and the edge, and the fringe is simply broadened in proportion to the distance l as l is increased. In the image diffraction serves, of course, to smooth out the sudden transition in intensity at the edge.

A somewhat similar, and frequently more extensive, smoothing takes place even for the electron concentration in a plane below the edge if the illuminating beam is not strictly parallel, but has a convergence α_c (Fig. 19·20). In this case some electrons transmitted by the free region

adjoining the edge enter the shadow of the specimen and the distribution becomes

$$J(y_o) = 1 - f\left(\frac{y_o}{l\alpha_c}\right) + k_2 f\left(\frac{y_o}{l\alpha}\right) \qquad 0 < y_o < l\alpha_c$$
$$= 1 + k_2 f\left(\frac{y_o}{l\alpha}\right) \qquad l\alpha_c < y_o < l\alpha \qquad [19\cdot61]$$

$$J(y_o) = k_2 - k_2 f\left(\frac{y_o}{l\alpha}\right) + f\left(\frac{y_o}{l\alpha_c}\right) \qquad -l\alpha_c < y_o < 0$$
$$= k_2 - k_2 f\left(\frac{y_o}{l\alpha}\right) \qquad -l\alpha < y_o < -l\alpha_c \qquad [19\cdot62]$$

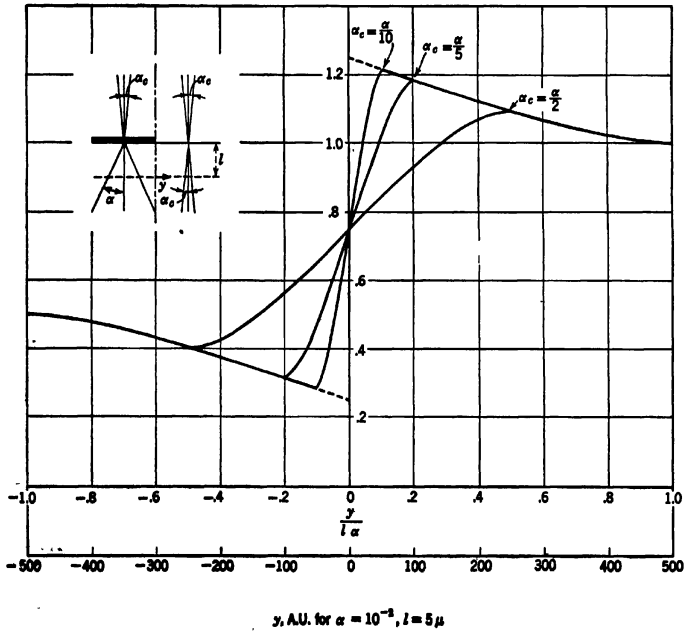


FIG. 19·20. Electron Distribution below Edge of Specimen Illuminated with a Convergent Beam of Electrons.

As shown in Fig. 19·20, the contrast is reduced progressively as the aperture of illumination α_c is increased. A scale in Angstrom units, calculated for the specific case $\alpha = 10^{-2}$ radian, $l = 5$ microns, gives an idea of the width of the fringe. It should be noted that, with parallel

illumination, the presence of a very thin film in the clear region, scattering the incident electrons into a very small solid angle, has an effect similar to a slight divergence of the illumination.



FIG. 19·21. Fresnel Diffraction Fringes and Defocusing Bands in the Image of a Filament of Artificial Rubber.

In the absence of a physical limiting aperture in the objective the outer edges of the bands cease to be sharply defined; apart from this the character of the fringes is little changed. Since here, for specimens not too thick, $k_2 = 1$, the contrast ratio at the edge becomes, for parallel illumination, $(2 + k_2) : k_2 = 3$. Contour fringes of the sort described, associated with the imperfect focusing of the specimen, are familiar to all users of the electron microscope.

All the contour effects discussed so far have the common property of leading to an intensity distribution which is antisymmetric about the geometric image of the specimen edge considered. This ceases to be true if the specimen has considerable thickness, so that an appreciable

number of image-forming electrons leave the specimen through a face substantially parallel to the optic axis of the microscope. Under these circumstances it may be shown that the bright band becomes more pronounced and narrower than the dark band, provided that the objective is focused either on the specimen or on a plane below the specimen.

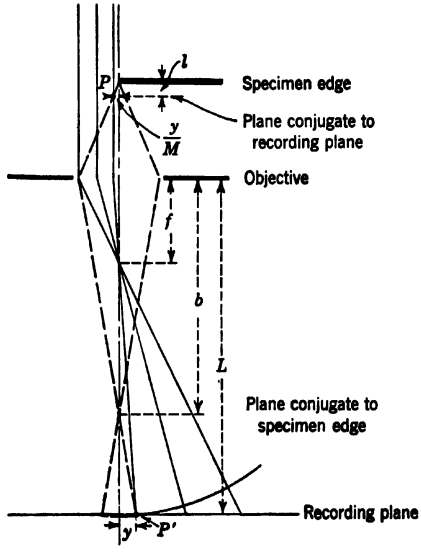


FIG. 19·22. The Origin of Fresnel Diffraction Fringes in the Out-of-Focus Image of a Specimen Edge.

A last contour phenomenon to be considered is observed in the form of fine fringes at the edges of thin specimens adjoining free space,[27] such as holes in thin object-supporting films. Figure 19·21 shows such fringes, in addition to the contour bands associated with imperfect focusing discussed above, at the edge of filaments of synthetic rubber. These fringes may be regarded as a consequence of the Fresnel diffraction of the coherent beam of electrons passing through the free region at the edge of the specimen.

Consider a uniform electron beam incident parallel to the axis. Let the specimen edge in question lie in the xz-plane, normal to the optic (z-) axis. Then beyond the specimen the electron beam may be regarded as made up of two waves: the plane wave which has passed

[27] See Hillier, reference 20, Boersch, reference 21*a*, and Ruska, reference 21*b*.

through the free region and a cylindrical wave with the edge of the specimen as source, advanced in phase by a quarter cycle with respect to the plane wave.[28] Let this combination of electron waves pass through the objective. If the objective is aberration-free (Fig. 19·22), the originally plane wave will converge in the lower focal plane of the objective and diverge as a spherical wave beyond it. Similarly, the objective will convert the cylindrical wave into a cylindrical wave converging on a line in the plane conjugate to the plane of the edge, a distance b from the image-side principal plane of the objective. Assume that the objective is focused not on the edge, but on a plane a distance l below it. The two electron waves will then interfere in the recording plane, a distance L from the image-side principal plane of the objective. Let Δ denote the difference in optical path for the cylindrical and the spherical wave from the plane of the edge to the point P'. P' is assumed to lie on the intersection of the yz-plane with the recording plane, a distance y from the optic axis. Then

$$\Delta = [y^2 + (L-b)^2]^{1/2} - (L-b) - [y^2 + (L-f)^2]^{1/2} + (L-f) \qquad [19\cdot63]$$

Solving for y,

$$\begin{aligned} y = \tfrac{1}{2}\{&8\Delta(b-f)(L-b)(L-f) \\ &+ 4\Delta^2[(b-f)^2 + (b-f)(L-f) - (L-f)^2] \\ &- 4\Delta^3(b-f) + \Delta^4\}^{1/2}/(b-f-\Delta) \end{aligned} \qquad [19\cdot64]$$

Since only low orders of interference, that is, path differences of a few wave lengths, are to be considered, it suffices to retain the term of lowest order in Δ:

$$y \cong \left[\frac{2\Delta(L-b)(L-f)}{b-f}\right]^{1/2} \qquad [19\cdot65]$$

Now, the magnification M in the recording plane is $(L-f)/f$ and, furthermore, by the elementary lens equation,

$$= \frac{f^2(L-b)}{(b-f)(L-f)} \qquad [19\cdot66]$$

Hence, Eq. 19·65 may be written

$$\frac{y}{M} \cong (2l\Delta)^{1/2} \qquad [19\cdot67]$$

The location of the maxima of interference, corresponding to

[28] See Wood, reference 22, Chapter 7.

$\Delta = (n + \frac{1}{4})\lambda$, is hence given by

$$\frac{y_{max}}{M} = [2l(n + \tfrac{1}{4})\lambda]^{1/2} \qquad [19{\cdot}68]$$

that of the minima by

$$\frac{y_{min}}{M} = [2l(n - \tfrac{1}{4})\lambda]^{1/2} \qquad [19{\cdot}69]$$

As might have been expected, these are the spacings of the maxima and minima in the Fresnel diffraction pattern in the plane below the edge which is conjugate to the recording plane; P' is simply the optical image

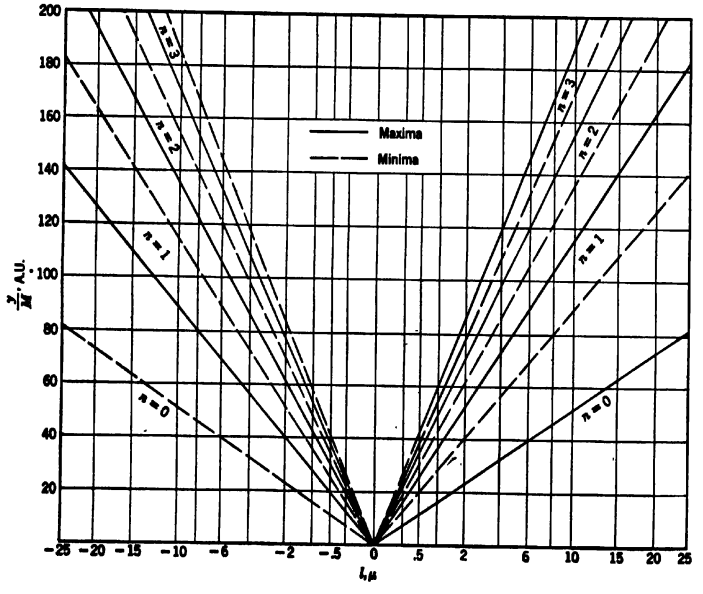


FIG. 19·23. Separation of the Maxima and Minima of Intensity of the Fresnel Diffraction Pattern from the Parallel Projection of the Specimen Edge.

of point P. It should be noted that the expressions 19·68 and 19·69 apply both for positive and negative values of l, that is, both if the recording plane is located below and if it is located above the true image of the edge.

The plot of y_{max}/M and y_{min}/M in Fig. 19·23 makes it appear that the image in Fig. 19·21 is out of focus by an amount of the order of 10 microns. This estimate may, however be somewhat in error since the

preceding considerations neglect the influence of spherical aberration. As indicated in Fig. 16·3, this aberration distorts the wave surfaces so that at the paraxial image of the edge the intersection of the cylindrical

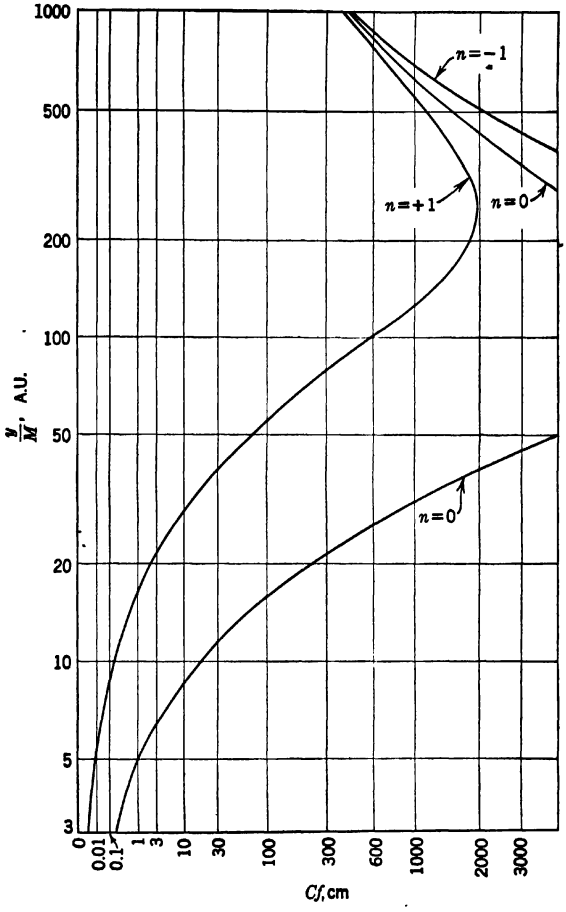


FIG. 19·24. Separation of Lowest-Order Maxima of Interference from Sharply Focused Image of Edge as Function of Spherical-Aberration Constant.

wave front with the yz-plane is not a point, but a curve. Hence a diffraction pattern is formed in the recording plane even if the edge is in exact (paraxial) focus. In this case the separations of the interference

maxima (bright fringes) from the edge are given, to the first order of approximation, by

$$\frac{y_{\max}}{M} = \left(\frac{f}{M}\right)^{1/4} (2\Delta)^{1/2} = \left(\frac{4}{3}\right)^{1/4} (Cf)^{1/4} \left(\Delta + \frac{\lambda}{4} + n\lambda\right)^{1/4} \qquad [19\cdot70]$$

those of the interference minima (dark fringes) by

$$\frac{y_{\max}}{M} = \left(\frac{f}{M}\right)^{1/4} (2\Delta)^{1/2} = \left(\frac{4}{3}\right)^{1/4} (Cf)^{1/4} \left(\Delta - \frac{\lambda}{4} + n\lambda\right)^{1/4} \qquad [19\cdot71]$$

Cf is here the spherical-aberration coefficient of the objective, Δ, the path difference for the originally plane wave between the ray to the point of reference and that to the image of the edge. Its determination involves the solution of a cubic equation. Figure 19·24 shows the position of the lowest-order maxima for the values $f = 0.3$ cm, $M = 130$, and $\lambda = 5.35 \cdot 10^{-10}$ cm (50,000 volts) as function of the aberration constant Cf. It is seen that the separations of the fringes in the image plane are of the order of 10 A.U. for values of Cf between 0.1 and 1, such as are commonly realized in modern electron-microscope objectives. They lie, thus, near the limit of resolution of present-day electron microscopes.

To summarize, both spherical and chromatic aberration give rise to contour phenomena. These phenomena are most prominent if the object field on one side of the contour has very low mass density and the illumination is nearly parallel. Under these circumstances a bright band appears on the light side of the image of the edge, a dark band on the dark side. A similar effect, independent of the lens aberrations, is observed with imperfect focusing of the objective, the width of the bands increasing in proportion with the degree of defocusing. In addition to these effects diffraction fringes may be observed parallel to the edge of the specimen. Their separation from the edge, again, increases as the specimen is thrown out of focus.

19·6. Resolution and Object Thickness. It is observed experimentally that with increasing thickness of the specimen the sharpness of the image decreases. With a narrow objective aperture this is due to two causes: First, electrons lose energy in inelastic collision within the object. As has been seen in section 19·2, this loss is far from uniform and results in a comparable spread of the energy of the electrons entering the objective. Owing to the chromatic aberration of the objective, electrons from any one point of the object are spread over a disk of diameter d_c:

$$\frac{d_c}{M} = k \cdot f \cdot \frac{\Delta\Phi}{\Phi} \cdot \alpha_m \qquad [19\cdot72]$$

where $e\Delta\Phi$ is the spread in the kinetic energy of the electrons and α_m the effective aperture of the objective. k is a constant characteristic of the objective; it is slightly smaller than unity for magnetic objectives.[29]

Second, electrons coming from the interior of the specimen are repeatedly deflected in their subsequent transit through the remainder of the specimen, so that they appear to come from points in the object plane differing from the actual object point. Consider, in Fig. 19·25, the electron rays originating in the object point O in the horizontal

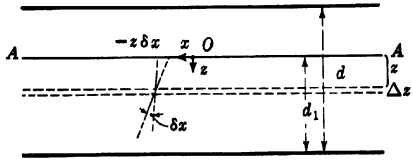


FIG. 19·25. Volume Scattering within the Specimen.

plane AA, a distance d_1 from the surface of the object facing the objective. If the layer d_1 is sufficiently thick, the electrons which have left O will be scattered in random fashion, so that, after leaving the object, they appear to come from a region about O in the plane AA, their apparent original distribution in this plane being given by the Gaussian error law:

$$J(r)r\cdot dr\cdot d\phi = e^{-r^2/(2P^2)}\frac{r\cdot dr\cdot d\phi}{2\pi P^2} \qquad [19\cdot73]$$

where P is the most probable radial deviation in apparent origin from O. As function of the distance x from a fixed line through O in the plane AA the distribution becomes

$$J_1(x)\,dx = \frac{e^{-x^2/(2P^2)}dx}{(2\pi P^2)^{1/2}}$$

and the mean value of x^2 is

$$\overline{x^2} = \int_{-\infty}^{\infty} x^2 J_1(x)\,dx = P^2$$

In order to determine P it is convenient to proceed as in the derivation of the most probable angle of scattering in multiple scattering (Eqs. 19·24–19·28). In any infinitesimal layer of the specimen the shift of the apparent origin of any electron coming from O is $-z\cdot\delta_x$, where z is the distance of the layer from the plane AA and $\delta_x = \delta\cos\phi$, the probability of the deflection within the layer being given, according to

[29] See section 17·6.

Eq. 19·17*a*, by

$$I(\delta)\delta d\delta d\phi = \left(\frac{Ze}{\Phi}\right)^2 N\,\Delta z\,\frac{d\delta\,d\phi}{\delta^3}$$

where δ may vary between $\delta_{\min}$ (determined by the effective radius R of the scattering atoms) and $\delta_{\max}$ (determined by the range of angles over which multiple scattering may be assumed). The mean square deflection of the apparent origin of the electrons due to scattering in the layer considered is

$$\overline{dx^2} = 2\int_{\delta\min}^{\delta_{\max}}\int_0^{\pi}(z\delta\cos\phi)^2 I(\delta)\delta d\delta d\phi = \pi\frac{Z^2e^2}{\Phi^2}Nz^2\,\Delta z\log\left(\frac{\delta_{\max}}{\delta_{\min}}\right)$$

The total mean square apparent deflection in the plane AA is given by the sum of these individual mean square deflections:

$$\overline{x^2} = P^2 = \Sigma\overline{dx^2} = \frac{\pi}{3}\left(\frac{Ze}{\Phi}\right)^2 Nd_1^3\log\left(\frac{\delta_{\max}}{\delta_{\min}}\right)$$

A comparison with Eq. 19·27 shows that this may be written, by substituting the proper values for $\delta_{\max}$ and $\delta_{\min}$,

$$P = d_1\cdot\frac{\Delta(d_1)}{3^{1/2}} = \frac{A^*d_1^{3/2}}{3^{1/2}} \qquad [19\cdot74a]$$

In this last equation $\Delta(d_1)$ is the most probable angle of deflection of the electrons in passing through the layer of thickness d_1, and A^* is the coefficient depending only on the operating voltage, the density of the specimen, and the atomic number of the latter:

$$A^* = \frac{8\cdot10^5}{\Phi}\,\frac{\Phi + 5\cdot11\cdot10^5}{\Phi + 10.22\cdot10^5}\cdot Z\left(\frac{\rho}{A}\right)^{1/2}$$

The diameter of the circle of confusion due to volume scattering may thus be put equal to

$$\frac{d_s}{M} = 2P = \frac{2}{3^{1/2}}A^*d_1^{3/2} \qquad [19\cdot74b]$$

It should be emphasized that d_1 is the thickness of the layer between the object structure observed and the surface nearest the objective, not the whole thickness of the specimen. For points on the surface of the specimen nearest the objective, the unsharpness due to volume scattering vanishes.

If the microscope is operated with a large physical objective aperture a third cause for a decrease in sharpness with an increase in object thickness is the increase in the effective aperture α_m resulting from the increase

in the mean scattering angle. This increases the effects of both spherical aberration and chromatic aberration.

For a specimen with $\rho = 1$, $Z = 6$, by Eq. 19·35*a*

$$\frac{\Delta\Phi}{\Phi} = \frac{1.4 \cdot 10^{11} d}{\Phi^2} \qquad [19\cdot75]$$

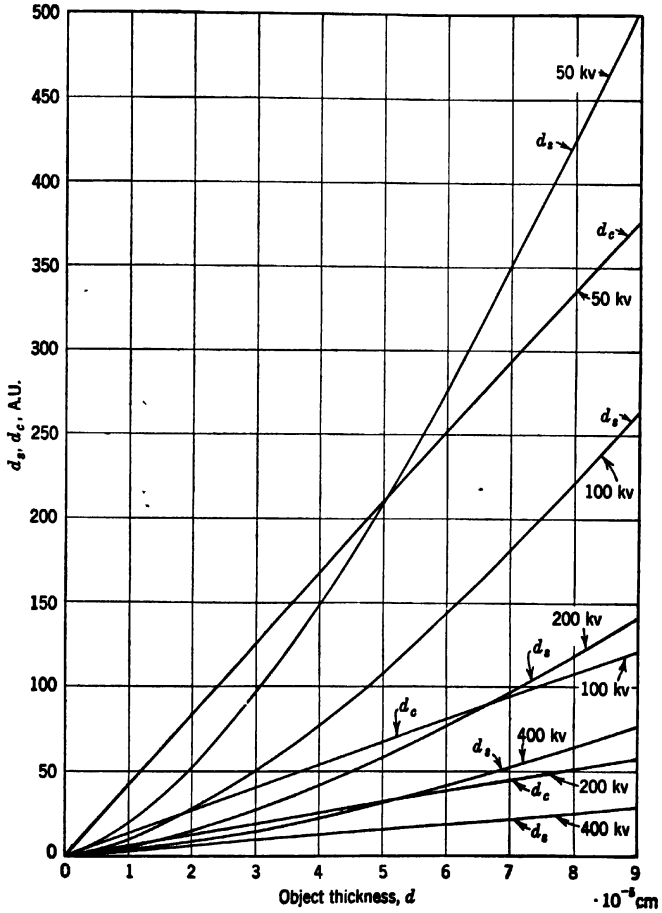


FIG. 19·26. Chromatic Aberration and Volume Scattering Effects.

and, furthermore, $A^* = 14$ centimeters$^{-\frac{1}{2}}$ for $\Phi = 50{,}000$ volts, being approximately inversely proportional to the applied voltage Φ. Thus the chromatic and the " volume scattering " defects become, for $f = 0.3$

centimeter and $k = 0.75$, Φ being set equal to 50,000 volts:

$$\frac{d_c}{M} = 12d\cdot\alpha_m \qquad \frac{d_s}{M} = 16d_1^{3/4}$$

For thin specimens this estimate of the chromatic defect is certainly too large in view of the smaller velocity losses of the electrons which are not scattered, or scattered only through small angles.

Figure 19·26 shows the variation of the two diameters of confusion, d_c and d_s, as a function of object thickness. In determining d_s, Eq. 19·74*b* for points at the center of the specimen ($d_1 = d/2$) was used. The aperture α_m has been set equal to $3.4\cdot10^{-3}$ for 50,000 volts and smaller values — keeping $f\alpha_m$ constant — for the higher voltages. Although at 50,000 volts the effect of volume scattering predominates only for values of the object thickness in excess of 0.5 micron, it becomes, for the narrow limiting apertures assumed, the more important of the two sources of unsharpness for normal specimen thicknesses at the higher voltages. For the large objective apertures normally employed, the increase in unsharpness with increasing specimen thickness is much more rapid than indicated by Fig. 19·26, owing to the increase in the effective objective aperture α_m. Thus, in addition to contrast considerations, the use of a narrow aperture becomes desirable in the study of very thick specimens from the standpoint of the best attainable resolution.

19·7. The Limit of Resolution of the Electron Microscope. The resolving power of any electron microscope is determined on the one hand by factors inherent in the character of the instrument and the properties of the object being examined and, on the other hand, by such as are incidental to the operation and construction of the unit considered. The latter — comprising the effects of voltage and field fluctuations due to imperfect stabilization, of inadequately shielded disturbing fields, and of misalignment and constructional defects — have been discussed in section 16·10 and do not concern the present consideration. The remainder are spherical aberration, diffraction, chromatic aberration due to velocity losses in the object, volume scattering in the object, and space-charge effects. Of these, again, space-charge effects play no role for the beam currents and voltages normally employed in the electron microscope, as may be deduced from Eqs. 16·208 and 16·210. The total current forming the image is normally of the order of 10^{-11} ampere.

With four fundamental defects, the question arises how they are to be added up to yield the least resolvable separation. The most reasonable simple procedure — applied by von Ardenne in the first attempt at a complete analysis of the resolution of the electron microscope[30] — con-

[30] See von Ardenne reference 23.

sists in setting this separation d_{tot} equal to the square root of the sum of the squares of the individual aberrations, that is, to add the aberrations like random errors:

$$d_{tot} = (d_{min}^2 + d_c^2 + d_s^2)^{1/2} \qquad [19\cdot76]$$

Here d_{min} is the least resolvable separation as determined by diffraction

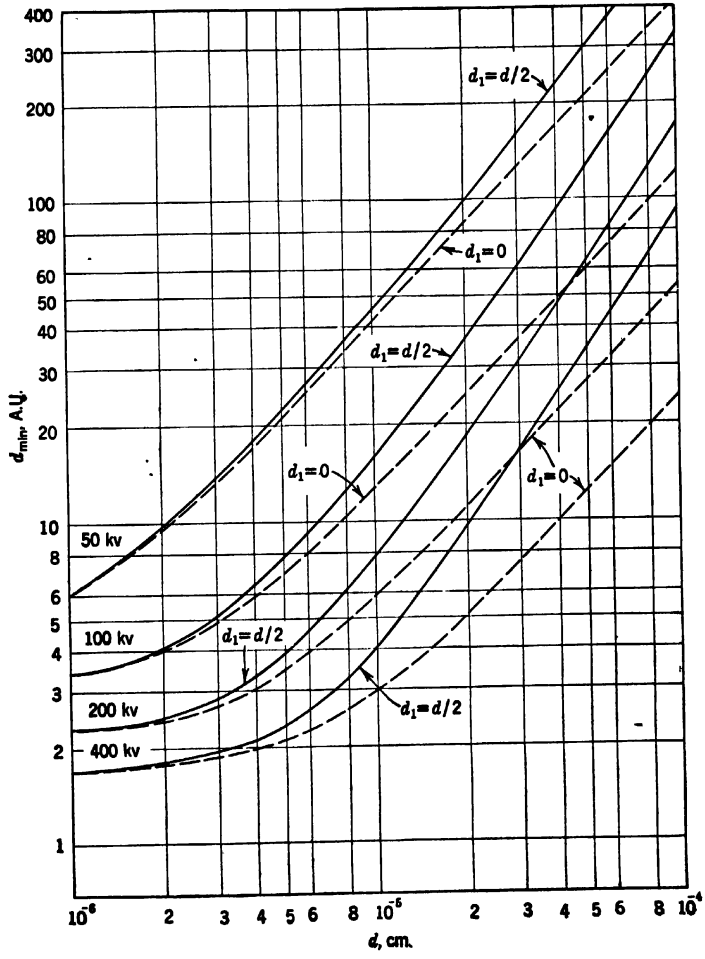


FIG. 19·27. Resolution of Ideal Objective as Function of Thickness ($C = 0.25$, $f = 0.1$ cm, $\alpha = \alpha_{opt}$).

and spherical aberration alone, given by Eqs. 19·50; d_c and d_s are given by Eqs. 19·72 and 19·74*b*, respectively.

Figure 19·27 shows the variation in d_{tot} with object mass thickness at different voltages, the minimum theoretical spherical aberration according to Rebsch,[31] $C = 0.25$, and $f = 0.1$ centimeter being assumed at all

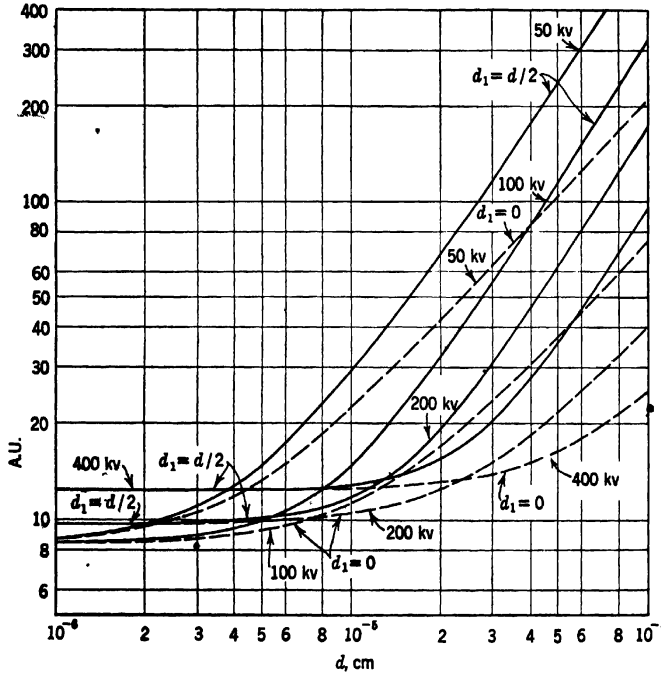


FIG. 19·28. Resolution of Objective as Function of Object Thickness ($C = 4$, $f = 0.1$ cm at 50,000 volts, $\alpha = \alpha_{opt}$ for $d \doteq 0$).

voltages. The objective aperture is assumed to be adjusted for optimum resolution for thin specimens. Under these circumstances there is a uniform improvement in resolution with voltage, this being greatest for large object thicknesses.

Figure 19·28 represents the variation in d_{tot} with object mass thickness under circumstances which are more readily realized in practice. Here the values $C = 4$, $f = 0.1$ centimeter have been assumed for 50,000 volts. For higher voltages the magnetic field is assumed to be

[31] See reference 18.

unaltered and the refractive power and spherical aberration to vary in accord with the second curve from the left of each set of curves in Fig. 18·1. In short, the iron of the lens is assumed to be saturated. Again, the aperture is adjusted for each voltage so as to attain best resolutions for thin specimens. Now the resolution becomes poorer with increasing voltage for light specimens, but improves for dense specimens. The continuous curves refer to objects examined at the center of the specimen ($d_1 = d/2$); the dotted curves refer to observation of the surface nearest to the objective ($d_1 = 0$).

All the curves in Figs. 19·27 and 19·28 refer to operation with small objective apertures. If a large objective aperture is employed, the value of d_{tot} will be only slightly increased for relatively thin specimens (of the order of 10^{-6} gram per square centimeter), but will be very considerably larger for dense specimens. The values for the chromatic aberration entering into the curves are somewhat too high since they do not take account of the fact that the average energy loss of the electrons accepted by the objective aperture is less than the average energy loss for all the electrons passing through the specimen. *It should be remarked, furthermore, that the measure of resolution given by these curves refers either to a transparent fine structure with the mass density indicated as abscissa on an opaque background or, conversely, to an opaque fine structure on a background of the mass density indicated. For the more complex objects normally encountered, in which neither the background nor the structure is opaque, these curves are not strictly applicable, but they may still serve as a guide.*

The most reliable method of experimentally determining the resolving power of the electron microscope is to locate two minute particles in the image field which appear barely separated and measure their distance between centers. Very thin heavy-metal films evaporated in vacuum, which, like gold, tend to break up into individual crystallites, form convenient objects for this purpose.[32] The values of the least resolvable distance so obtained may be regarded as conservative.

The results of some measurements carried out in this manner by R. F. Baker of the RCA Laboratories are collected in Fig. 19·29. Two series of measurements were made: one with a physical objective aperture of diameter D, $D/(2f) = 9{\cdot}10^{-3}$ radian; the other without any physical aperture. In each case the aperture of the illumination was varied from $3{\cdot}10^{-4}$ radian to $3.8{\cdot}10^{-3}$ radian by inserting condenser apertures of various sizes. The striking feature of the results is the strong dependence of the resolution on the aperture of illumination, even for relatively low values of the latter. This may conceivably be ascribed

[32] See von Ardenne, reference 24.

to imperfect alignment of the electron beam and the objective, causing the effective angle of illumination to be larger than that plotted as abscissa. The least resolvable distances are seen to be throughout about twice as great for operation without an aperture as for operation with the aperture. The reduced contrast of the gold grains in the former case,

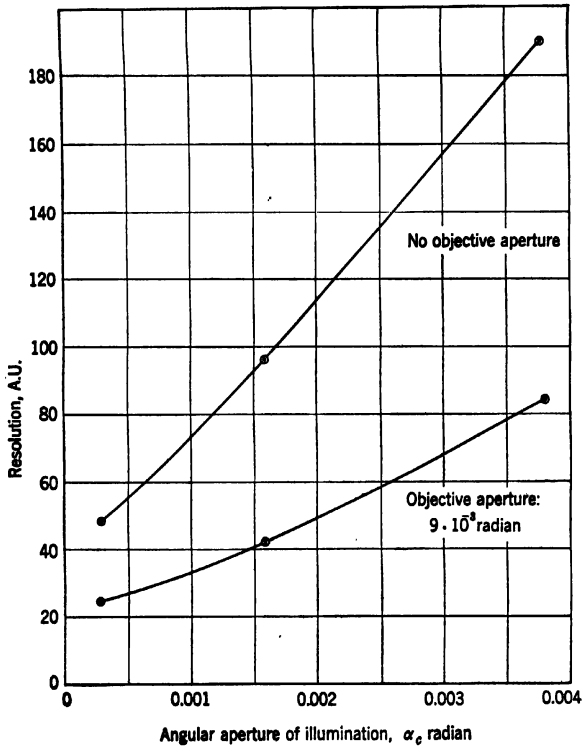


FIG. 19·29. Measured Resolution of an Objective as Function of the Angular Aperture of the Illuminating Pencils, α_c (R. F. Baker.)

rendering the recognition of just-separated grains more difficult, may have enhanced the actual difference slightly. It is to be assumed that here, as in other cases, part of the difference between the theoretically expected and the actually measured values of the resolution must be ascribed to disturbing experimental factors which never can be completely eliminated. It is evident from Fig. 19·20 that imperfect focusing of the specimen could have been responsible for the observed dependence of the resolution on the aperture of illumination.

19·8. Limits of the Recognition of Small Objects; Dark-Field Operation. A very small particle illuminated by a uniform beam of electrons appears, in the image, as a small "dip" in the intensity (see, for example, Fig. 19·14). The area of this "dip" is approximately equal to $\pi d_{\min}^2 \cdot M^2$, where $d_{\min}$ is the lea- resolvable separation, assuming that the diameter of the particle is substantially less than $d_{\min}$. The mean intensity within this dip compared to the intensity of the background will be given simply by

$$1 - \frac{\sigma}{\pi d_{\min}^2}$$

where σ is the scattering cross section of the particle, whereas that of the center of the dip may be estimated to be (assuming an intensity distribution corresponding to that in the Airy disk)

$$1 - \frac{\sigma}{d_{\min}^2} \qquad [19\cdot77a]$$

For a heavy atom, practically all the scattering being nuclear, σ is given, according to Eq. 19·20*a*, by

$$\sigma = 0.56\lambda^2 Z^{4/3}$$

In order that the particle may be recognized, the maximum difference in intensity of the dip and of its background must exceed a certain percentage of the background intensity. The magnitude of this percentage depends on the uniformity, the contrast properties, and the grain size of the electron indicator — for example, the photographic plate — used. If a difference of intensity of 2 per cent can still be recognized, a single atom can be seen or recorded in the image provided that

$$Z > 0.082 \left(\frac{d_{\min}}{\lambda}\right)^{3/2}$$

Thus, for $\lambda = 5.35 \cdot 10^{-10}$ (50,000 volts) and $d_{\min} = 15 \cdot 10^{-8}$ centimeter, $Z > 380$. This would suggest that even for as heavy an element as gold at least eight atoms would have to make up the particle in order to render it visible in present-day electron microscopes.

The above calculation can be regarded as substantially correct only if the resolution is determined by geometrical aberrations rather than diffraction phenomena, or if the incident electron beam has been rendered incoherent, for example, by previous passage through a substantial layer of matter. Otherwise, as Schiff[33] has shown, account must be taken of the interference of the elastically scattered wave with the incident plane

[33] See reference 25.

wave. Schiff proceeds in the following manner in determining the lightest single atom that may be recognized with an ideal (that is, aberration-free) electron microscope. In view of the rapid falling off of the elastic scattering beyond the limiting angle δ_1, Eq. 19·20c, in accord with the Rutherford law, the angle δ_1 may be considered as determining the maximum effective aperture of the objective, so that in the image plane the scattered intensity is spread over an Airy disk of radius a:

$$\frac{a}{M} = \frac{0.61\lambda}{\delta_1}$$

If A is the amplitude of the incident wave, $A^2\sigma_n$ is the total elastically scattered current. The root-mean-square amplitude of the scattered wave within the Airy disk, being equal to the square root of the mean scattered current per unit area, is $(A^2\sigma_n/[\pi a^2])^{1/2}$, and the contrast g becomes equal to

$$g = 1 - \frac{\left\{\frac{A}{M} \pm \left(\frac{A^2\sigma_n}{\pi a^2}\right)^{1/2}\right\}^2}{\left(\frac{A}{M}\right)^2} = 2\left(\frac{\sigma_n}{\pi a^2}\frac{}{M^2}\right)^{1/2} \qquad [19\cdot77b]$$

By substituting the values for σ_n and δ_1 from Eqs. 19·20b and 19·20c,

$$g = \frac{0.008Zc}{v}$$

where v is the velocity of the electrons and c is that of light. For $g = 0.10$, $Z < 6$ for 50,000 volt electrons, whereas for the less conservative value $g = 0.02$ (2 per cent) for the least observable contrast all atoms heavier than hydrogen can be observed.

These results cannot be applied directly to practical present-day microscopes operating at the usual voltages, since for these spherical aberration exerts a predominating influence at the angles δ_1 corresponding to voltages and atomic numbers which give adequate contrast. If, however, electron microscopes are built which have spherical aberration coefficients appreciably less than 1 centimeter even when operated at very high voltages, this is no longer the case. Thus for an accelerating voltage of one million volts and an atomic number 8, $\delta_1 = 4\cdot10^{-3}$ and $g = 0.06$. The recognition of individual atoms is thus not an impossibility. A serious practical difficulty consists, of course, in immobilizing the atoms without destroying the coherence of the scattered electrons and without swamping the effect of the atom considered with the scattering from the atoms of the supporting medium.

Fundamentally, as Schiff also points out and as is familiar from the

practice of light ultramicroscopy, observation in dark field offers a way of circumventing the contrast difficulties. A dark field is most readily obtained either by placing an annular aperture in the condenser lens (Fig. 19·30*a*), making the illuminating angle larger than the angle accepted by the objective, or by deflecting the illuminating beam just enough to let it miss the objective aperture (Fig. 19·30*b*). The latter

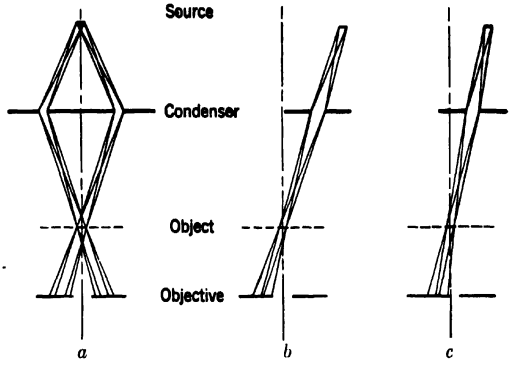


FIG. 19·30. Dark-Field Illumination (*a*) with Ring Aperture in the Condenser, (*b*) by Deflection of Illuminating Beam, and (*c*) by Deflection of Illuminating Beam and Partial Masking of Objective Aperture.

method is usually simpler. It is clear that under these circumstances only the scattered radiation reaches the image, so that, with sufficiently long exposure, an isolated particle can be recorded, no matter what its size or scattering power. Practically, again, the necessity of supporting the particle, adding the scattering of the supporting medium to that of the particle, reduces the value of the method. It may be expected to be advantageous, however, in the observation of fiberlike structure, such as long molecular chains, which need be supported only at their ends.

In general the resolution obtained in dark-field pictures is lower than in bright-field pictures. The reason for this is not definitely known. With a large physical aperture the greater spherical aberration at the edge, being given by

$$\Delta r = 3Cf\alpha^2 d\alpha \qquad [19·78]$$

for $d\alpha \ll \alpha$, results in poorer definition. Here α is the aperture angle corresponding to the physical aperture of the objective and $d\alpha$ the effective aperture angle of the imaging beam. This difficulty could be avoided by masking off just half the physical aperture of the objective (Fig. 19·30*c*). Even under these circumstances, however, the influence of spherical aberration would be greater than for bright-field operation,

since the intensity of the imaging beam falls off more gradually in a direction normal to the deflection.

Although, in principle, it would be possible to gain in resolution by providing a very thin annular aperture in the objective corresponding to a large value of α, this becomes impractical because of the exceedingly small dimensions involved. Thus, if the resolving power were to be increased by a factor of 3 ($d_{\min}$ = 5 A.U. in place of 13 A.U. with Cf = 2 centimeters),

$$5{\cdot}10^{-8} \cong 3Cf\alpha^2 d\alpha \quad \text{or} \quad d\alpha = 10^{-4} \text{ radian}$$

Even for an objective with a focal length of 1 centimeter, the width of the annular aperture would have to be only 1 micron. In general, this width is proportional to the cube of the least distance to be resolved.

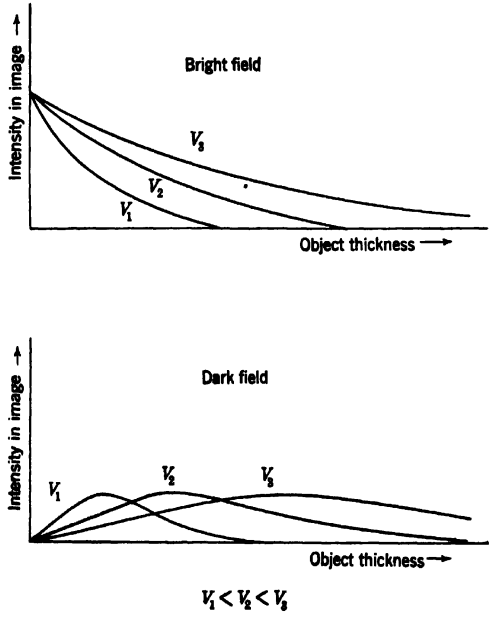
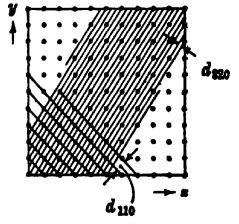


FIG. 19-31. Variation of Image Brightness with Object Thickness for Bright and Dark Field at Different Voltages (Schematic).

Some reference should be made to the variation of brightness with object thickness in dark-field images. Completely unoccupied regions of the object field appear, obviously, black. For portions of the specimen which are so thin that for them single scattering takes place, there is

a linear relation between the brightness and the mass thickness. This range is larger in proportion as the decentering of the primary beam is greater. With the onset of multiple scattering the increase in brightness with mass thickness becomes slower. At a point where the most probable angle of scattering is equal to $2^{-\frac{1}{2}}$ times the angle between the direction of the primary beam and the direction of observation, the brightness of the image attains a maximum. Beyond this point the intensity falls off again. An increase in the operating voltage has the effect of increasing the thicknesses for which a given intensity is attained, the increase varying linearly with the voltage for small thicknesses. Figure 19·31 shows qualitatively the variation of brightness with mass thickness at different voltages in bright field and in dark field.[34]

19·9. Crystalline Diffraction Effects. Allusion has been made to the fact that deviations from a simple relationship between intensity of the image and mass density of the object occur if the atoms of the specimen are arranged in a periodic lattice as is the case for crystalline materials. For certain orientations of the crystal relative to the incident beam the electrons are strongly reflected, reducing the intensity of the incident beam correspondingly. Thus for such orientations even a very thin crystal will appear dark in the image.



The condition for the existence of such selective reflections is known as the Bragg law. The atoms or ions making up any crystal may be regarded as lying in any one of a large number of sets of equidistant parallel planes. The separation of two successive planes is denoted by $d_{h,k,l}$, h, k, and l being integers — the *Miller indices* — characterizing the set of planes considered (Fig. 19·32). $d_{h,k,l}$, normally abbreviated as d, is a simple function of h, k, l, which for a cubic crystal is

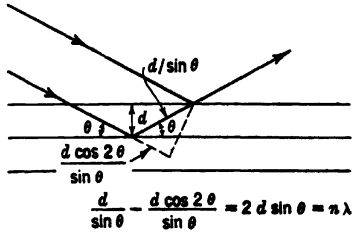
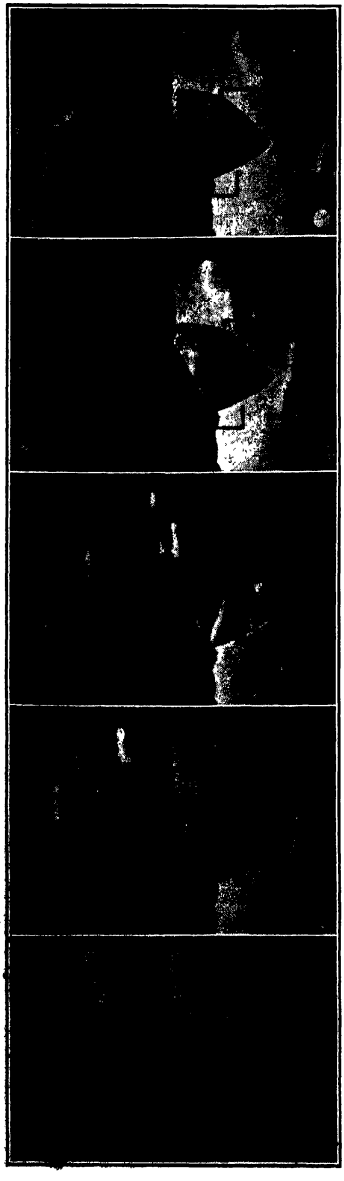


FIG. 19·32. The Condition for Reflection at a Crystal Lattice.

$$d = \frac{a}{(h^2 + k^2 + l^2)^{\frac{1}{2}}} \qquad [19·79]$$

[34] See v. Borries and Ruska, reference 26.



Here a is the least distance between two atoms or ions (a side of the elementary cell or cube) and a/h, a/k, a/l are the intercepts on the three sides of the elementary cube of the plane of the set nearest to the atom treated as origin. If θ is the angle between the incident beam and a particular set of planes with separation d, specular reflection will take place at these planes if Bragg's law

$$n\lambda = 2d \sin \theta \qquad [19{\cdot}80]$$

is fulfilled. It follows from the geometry of the figure that this is the condition when the path difference for rays reflected at successive planes is just $n\lambda$, causing them to reinforce each other. Figure 19·33[35] shows dark bands on the characteristically shaped thin lamellae of aluminum oxide monohydrate (diaspore or boehmite), which are due to such crystalline reflection. The condition 19·80 must be fulfilled so exactly that the distortion of the crystal or the divergence of the illuminating beam prevents reflection from taking place over the entire area of the crystal.

In addition to the dark bands, similar bright bands are observed (Fig. 19·33) which move across the dark bands as the focus of the

FIG. 19·33. The Effect of Crystalline Reflections on Images of Aluminum Oxide Monohydrate Scales for Different Focusing Currents. (Courtesy of *Phys. Rev.*, reference 27.)

[35] See Hillier and Baker, reference 27.

objective is changed. These appear only when the physical aperture of the objective is large. They are formed by the reflected electrons (Fig. 19·34). Owing to the large spherical aberration of the lens, they coincide with the dark bands formed by the practically undeviated beam for objective current values different from those required for optimum focus of the principal image. Furthermore, in view of the large angle of incidence on the lens the image due to the reflected electrons is deflected considerably by changes in lens focus which affect the main image only slightly. More complex band structures, attributed to multiple reflections of electrons, have been observed by Heidenreich[36] in magnesium oxide crystals.

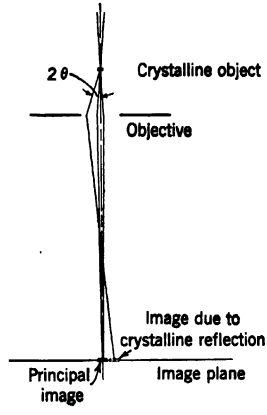


FIG. 19·34. Origin of the Bright Bands and Spots Due to Crystalline Reflection.

19·10. Limits of Resolution of Other Types of Electron Microscopes. It can readily be shown that the ultimate limits of resolving power of other electron microscopes, such as the scanning microscope, the shadow microscope, the x-ray shadow microscope, the point projector microscope, and the emission microscope, either correspond to those of the standard transmission instrument or are less favorable. The enumerated devices will now be discussed individually.

a. The Scanning Microscope. In the scanning microscope a fine electron probe scans the specimen. The current of the secondary or transmitted electrons leaving the specimen as the result of the impact of the primary electrons of the probe controls, in some suitable manner, the intensity of an image scanned in synchronism. Variations in the emitting or transmitting power of the object can thus appear in the image only if they extend over a distance of the same order or larger than the probe diameter. The dimension of the probe thus determines the ultimate resolving power of the instrument. Since diffraction at the aperture of the final reducing lens (Fig. 19·35) gives rise to an Airy diffraction disk with a separation of the maximum and the first minimum equal to

$$d_{\text{diff}} = \frac{0.61\lambda}{\alpha}$$

where α is the angular aperture of the probe and, since, furthermore, the

[36] See reference 28.

spherical aberration of the same lens forms a circle of confusion of radius

$$d_{\text{sph}} = Cf\alpha^3$$

where C is the spherical aberration coefficient of the final lens and f its focal length, the ultimate limit of resolution $d_{\min}$ is given by exactly the same formula as for the standard electron microscope (Eqs. 19·47 and 19·50*b*).

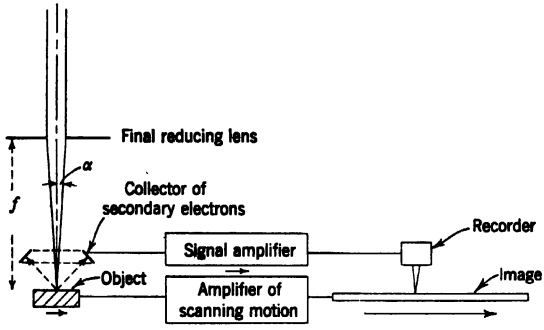


FIG. 19·35. Image Formation in the Scanning Microscope (Schematic.)

In one respect, as has been pointed out by von Ardenne,[37] the scanning microscope has a special advantage. The resolution is independent of the velocity distribution of the electrons which leave the object and serve to determine the image contrasts. For a transmission scanning microscope this signifies that sharp images can be obtained of objects lying at the surface facing the final lens of relatively thick specimens. The importance of this advantage is here reduced by the poor contrast conditions of such objects and the confusion resulting from the superposition of unsharp images of inner structure of the specimen. However, the same characteristic makes possible the observation, with the scanning microscope, of the surfaces of opaque specimens with a resolution of the same order as that obtained with transparent objects in the standard electron microscope. The relatively low and inhomogeneous velocities of the secondary electrons providing the signal current have no effect on the image quality.

b. The Shadow Microscope.[38] In the shadow microscope an electron probe is formed in the same manner as in the scanning microscope. The (transparent) object, however, is not placed at the probe, but a small distance beyond it, and the probe, being maintained in stationary posi-

[37] See reference 29.

[38] See Boersch, reference 30, and Scherzer, reference 15.

tion, projects a greatly enlarged shadow image on a screen. The magnification of this image is equal to the ratio of the distance of the screen from the probe, a, and that, b, of the object from the probe (Fig. 19·36). The contrast conditions resemble those of the standard electron microscope operating with a large objective aperture. The physical aperture in the latter case corresponds to the field-limiting aperture in the shadow microscope.

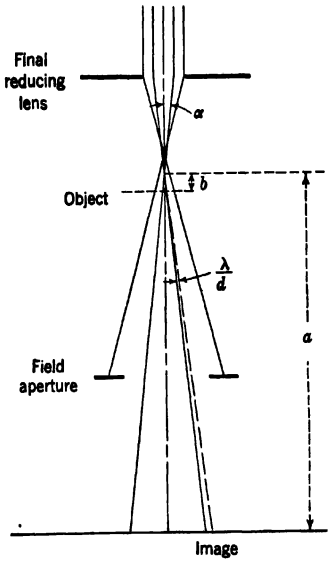


FIG. 19·36. The Shadow Microscope.

As for the instruments previously discussed, the effective resolution of the shadow microscope is limited by diffraction and by the spherical aberration of the final lens. The diffraction here to be considered takes place at the object. For a periodic structure of grating spacing d the angle of diffraction corresponding to the first maximum is λ/d. In order that the image of this structure be resolved, the resulting diffraction figure on the screen must be narrower than the separation of the elements of the structure on the image, ad/b. Since the diffraction figure has a width $\lambda a/d$, the condition for the resolution of the structure becomes

$$d \geqq (\lambda \cdot b)^{1/2} \qquad [19\cdot81]$$

For a simple edge the same condition holds if the least resolvable distance is here interpreted as the separation of the geometrical image of the edge and the first maximum of the Fresnel diffraction pattern.[39] In the shadow microscope the spherical aberration of the final reducing lens does not lead to unsharpness of the image, but causes it to be distorted, the effective position of the probe moving away from the specimen with increasing angle α. Thus

$$b = b_o + C f \alpha^2$$

If, owing to this effect, b is doubled or the magnification is halved for a spacing in the object d corresponding to an angle $\alpha = d/b_o$, the structure

[39] See Houstoun, reference 31, pp. 164 ff.

becomes unrecognizable. Thus b_o must be chosen at least so large that

$$b_o = Cf\left(\frac{d}{b_o}\right)^2$$

Substituting the corresponding value for b_o for b in Eq. 19·81 makes the limit of resolution

$$d_{\min} = \lambda^{3/4}(Cf)^{1/4} \qquad [19\cdot82]$$

which is equivalent to Eq. 19·50*b*, applying for the standard electron microscope.

c. The X-ray Shadow Microscope. Here the electron probe, formed as in the two preceding types of microscope, impinges on a thin anticathode. The x-rays generated at the point of impact project an enlarged image of a closely situated object (Fig. 19·37) on a screen. The ultimate limit of resolution is given here, as in the case of the scanning microscope, by the size of the probe — that is, is identical with that for all previously discussed instruments. However, unless the distance b is chosen exceedingly small (of the order of 10^{-6} centimeter), the diffraction of the x-rays at the object limits the performance of the instrument in accordance with Eq. 19·81, the electron wave length λ being replaced by the x-ray wave length λ_r.[40] Since the maximum of intensity of the x-radiation lies approximately at the frequency ν_r given by

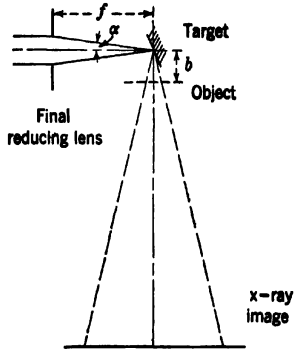


FIG. 19·37. The X-ray Shadow Microscope.

$$h\nu_r = \frac{hc}{\lambda_r} = \frac{1}{2}\frac{mv^2}{2} = \frac{h^2}{4m\lambda^2}$$

this added restriction on the resolution may be written in terms of the electron wave length λ:

$$d \geqq 2\lambda\left(\frac{b}{\lambda_c}\right)^{1/2} \qquad \lambda_c = 2.3\cdot10^{-10}\text{ cm} \qquad [19\cdot83]$$

d. The Point Projector Microscope.[41] The point projector microscope is the only transmission microscope which makes use of no electron

[40] See Scherzer, reference 15.

[41] See Morton and Ramberg, reference 32.

lenses. It consists of a fine point formed by etching at the end of a wire (usually of tungsten) between which and an anode, which may be the object holder itself, a high voltage is applied (Fig. 19·38). As in the shadow microscope, the source projects an enlarged image of the object placed in its path.

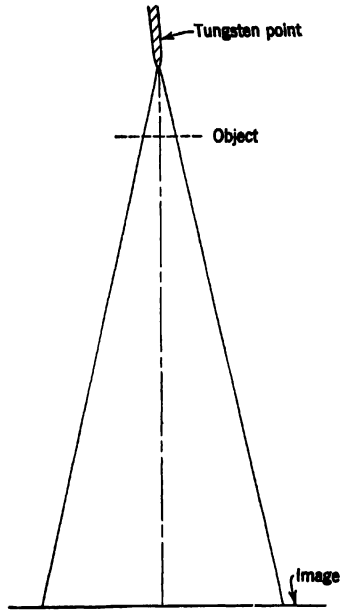


FIG. 19·38. The Point Projector Microscope.

The resolution of this instrument is limited by two factors: first, Fresnel diffraction at the object, which, in accordance with Eq. 19·81, determines the maximum distance b between source and object which can be used if a given desired resolution is to be attained; second, the unsharpness produced by the initial velocities of the electrons emitted by the point.

To determine the last effect it is convenient to treat the field as though it were spherically symmetrical, the radius of the point being designated by r_o and that of the anode at potential V by $R \gg r_o$. As Ruska[42] has shown, the electrons leaving the surface of the point describe hyperbolas in this field. In particular, if eV_r is the initial energy of an electron emitted tangent to the surface, the electron reaches the anode with an angular displacement ϕ_a (Fig. 19·39),

$$\phi_a = 2\left(1 - \frac{r_o}{R}\right)\left(\frac{V_r}{V}\right)^{1/2} \qquad [19·84]$$

and an inclination to the radius through the point of incidence ψ_a,

$$\psi_a = \frac{r_o}{R}\left(\frac{V_r}{V}\right)^{1/2} \qquad [19·85]$$

The only assumption is here that V_r is much smaller than V.[43]

[42] See reference 33.

[43] The expressions in Eqs. 19·84 and 19·85 become exact if V is replaced by $V + 2V_r(1 - r_o/R)$.

The figure shows that the diameter of the circle of confusion resulting from electrons with lateral velocities $(2eV_r/m)^{1/2}$ is

$$d = \frac{2}{M}(L - R)\psi_a = 2r_o\left(\frac{V_r}{V}\right)^{1/2} \qquad [19\cdot86]$$

if the magnification M is large or $L \gg R$. For a given point the voltage V must be so adjusted that the cold-emission current is adequate and yet not large enough to result in the speedy destruction of the point. For tungsten, from many points of view the most favorable material for the cold emitter, the requisite surface field is about $3\cdot10^7$ volts per centimeter. Since the surface field is given by

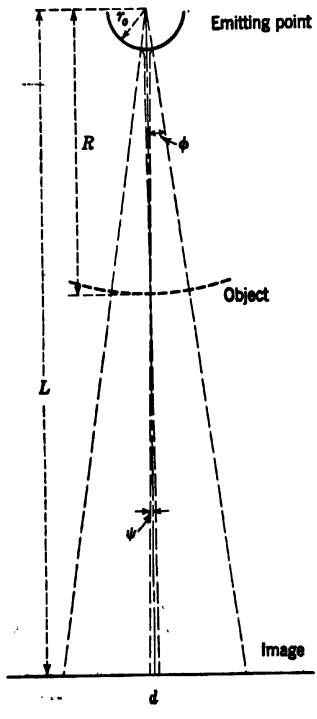


FIG. 19·39. Effect of Initial Velocities of Electrons in Point Projector Microscope.

$$\Phi_o' = \frac{VR}{r_o(R - r_o)} = 3\cdot10^7 \text{ volts per cm} \qquad [19\cdot87]$$

Equation 19·86 may be written in the form

$$d = \frac{6.7\cdot10^{-8}}{1 - \dfrac{r_o}{R}}(VV_r)^{1/2} \qquad [19\cdot88]$$

Detailed calculations for hyperboloidal points and plane anodes indicate that Eq. 19·88 applies also to these. r_o here becomes the radius of curvature at the vertex of the hyperboloid; R the distance between the center of curvature and the plane anode.

Calculations by Richter[44] indicate that the mean lateral kinetic energy of the field electrons from tungsten, for a surface field of $3\cdot10^7$ volts per centimeter, is $eV_r = 0.09$ electron-volt. Hence,

$$d = \frac{2\cdot10^{-8}}{1 - \dfrac{r_o}{R}}V^{1/2} \qquad [19\cdot89]$$

[44] See reference 34.

If, to this, the unsharpness due to Fresnel diffraction (Eq. 19·81) is added, with the wave length expressed in terms of the applied voltage, the resolution of the point projector microscope becomes

$$d_{\min} = \frac{2\cdot 10^{-8}}{1 - \dfrac{r_o}{R}} V^{1/2} + 3.51\cdot 10^{-4} \frac{R^{1/2}}{V^{1/4}} \qquad [19\cdot 90]$$

For given V this quantity becomes a minimum for a particular value of r_o/R. Thus, for $V = 10{,}000$ volts, r_o/R should be 0.2; for $V = 100$ volts, it should be 0.3. There is a continuous improvement in resolution as the voltage (and hence both the point radius and the distance between the point and the object) is decreased. If 100 volts is accepted as the lowest voltage at which a stable discharge can be maintained, an optimum resolution of 73 A.U. is obtained with a point with 475 A.U. radius. At an operating voltage of 10,000 volts a resolution of 410 A.U. may be obtained with a point radius of 4 microns.

e. The Emission Microscope. The term *emission microscope* includes all microscopes in which the object itself is the cathode and emits electrons of velocities ranging from zero to some value small compared to the velocity in image space.

A very special case is the point field-emission microscope of E. W. Müller[45] for studying the emission properties of point cathodes. If the point is treated as a sphere of radius r_o surrounded by a spherical anode of radius R, the diameter of the circle of confusion arising from the initial velocities becomes, by Eq. 19·84 and Fig. 19·39,

$$l = \frac{2R}{M}\phi_a = 4r_o\left(\frac{V_r}{V}\right)^{1/2} \qquad [19\cdot 91]$$

It is here permissible to neglect r_o relative to R. If it is assumed, as for the point projector microscope, that the field at the surface of the point is $3\cdot 10^7$ volts per centimeter and that the mean lateral energy of emission is 0.09 electron-volt, the expression in Eq. 19·91 becomes

$$d = 4\cdot 10^{-8} V^{1/2} \text{ cm} \qquad [19\cdot 92]$$

This relation applies also approximately for a hyperboloidal point and a plane anode, a condition more readily realized in practice. It is found that for such a configuration the numerical coefficient in Eq. 19·92 decreases gradually as the magnification of the instrument is increased. For a magnification of a million it has dropped to about $1.3\cdot 10^{-8}$.

The behavior of the hyperboloidal point emission microscope differs in

[45] See reference 35.

two other respects from the idealized spherically symmetrical cold-emission microscope. First, the magnification of the instrument is $k \cdot R/r_o$, where r_o is the radius of curvature of the point and R the point-to-anode distance, rather than simply R/r_o. k is a constant which decreases gradually with increasing magnification. It is approximately 0.5 for a magnification of a million.

Second, the voltage required, for given r_o and R, to produce the necessary surface field is considerably greater for the hyperboloidal point microscope. It may be obtained from the relation[46]

$$\Phi_o' = \frac{2V}{r_o \log \frac{4R}{r_o}} \qquad [19\cdot93]$$

For example, to achieve a magnification of a million with an image distance $R = 10$ centimeters the hyperboloidal point would have to have a radius of about 0.05 micron and the operating voltage would have to be 1180 volts. The same result would be attained in the spherically symmetrical configuration with $r_o = 0.1$ micron and $V = 300$ volts. The diameter of the circle of confusion, as given by Eq. 19·92 with the numerical coefficient appropriately modified, is, however, less for the hyperboloidal point microscope. It is about 45 A.U. in the above example. For an operating voltage of 100 volts it drops to 12 A.U. The actual least-resolvable distance is, however, larger than this. An application of the uncertainty principle, as on page 695, would lead to an apparent diffusion of an object point by $h/(2mv_r) = \lambda_r/2$, where v_r is the mean radial velocity of the electron at the object and λ_r the corresponding de Broglie wave length. For $V_r = 0.09$ electron-volt this leads to an ultimate limit of resolution of the order of 20 A.U.

Emission microscopes with flat objects will be the subject of the remainder of the section. They include thermionic, photoemission, and secondary-emission instruments, the object being heated, illuminated, or bombarded with primary electrons to cause it to emit the electrons which form the image.

General limits of resolution can readily be determined for this class of electron microscopes, since the spherical and chromatic aberration introduced by the accelerating field at the surface of the cathode is normally much greater than that arising from any other part of the imaging system. Although the final formulas obtained for the resolution are independent of the arrangement of the system, it is convenient to regard it as made up of an initial uniform accelerating field and a unipotential

[46] See Eyring, Makeown, and Millikan, reference 36.

objective lens (Fig. 19·40). Let the beam aperture of the electrons leaving the object be α_o, that of the electrons entering the objective lens, α. Furthermore, let the range of initial kinetic energies be from 0 to $e\Phi_o$ and let the kinetic energy of the electrons entering the lens be $e\Phi$. Then, by Eq. 16·151, the apparent point of divergence on the axis of electrons leaving with an axial initial velocity $(2e\Phi_z/m)^{1/2}$ is given by

$$\frac{1}{z^*} = \frac{r'}{r} = \frac{1}{2z}\left(1 + \left[\frac{\Phi_z}{\Phi_o' z}\right]^{1/2} \cdots\right) \quad [19\cdot94]$$

and the separation of the apparent points of divergence for electrons with this initial axial velocity and with zero initial axial velocity is

$$\Delta z = (2z)^2 \Delta\left(\frac{1}{z^*}\right) = 2\left(\frac{z\Phi_z}{\Phi_o'}\right)^{1/2} \quad [19\cdot95]$$

If $(2e\Phi_r/m)^{1/2}$ is the initial lateral component of velocity, there will thus be a circle of confusion in the virtual object plane of the objective lens with a radius

$$\Delta r = \left(\frac{\Phi_r}{\Phi_o' z}\right)^{1/2} \Delta z \doteq \frac{2(\Phi_r \Phi_z)^{1/2}}{\Phi_o'} \quad [19\cdot96]$$

Two cases must now be considered: the presence of a limiting aperture in the objective admitting electron pencils with a maximum aperture

$$\alpha = \left(\frac{\Phi_r}{\Phi}\right)^{1/2} = \alpha_o \left(\frac{\Phi_o}{\Phi}\right)^{1/2}$$

and the absence of a limiting aperture. In the first case it is appropriate to put $\Phi_r = \Phi_o \sin^2 \alpha_o = \Phi_o \alpha_o^2$ and $\Phi_z = \Phi_o \cos^2 \alpha_o = \Phi_o$, so that

$$\Delta r = \frac{2\alpha_o \Phi_o}{\Phi_o'} \quad [19\cdot97]$$

If the lens is focused halfway between the virtual object plane for rays with zero and with maximum initial axial velocity the diameter of the circle of confusion, rather than the radius, is given by the expression in Eq. 19·97.

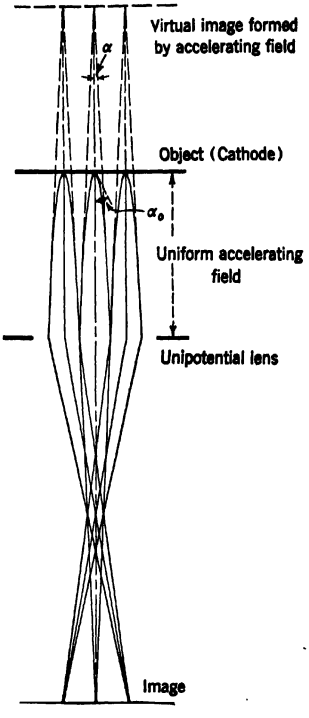


FIG. 19·40. Schematic Representation of an Emission Microscope.

The limit of resolution and the optimum lens aperture are determined by minimizing the sum $2\alpha_o\Phi_o/\Phi_o' + 0.6\lambda/\alpha_o$. The second term represents the radius of the diffraction disk in the absence of other aberrations. The values of α_o and $d_{\min}$ thus obtained are:

$$\alpha_o = 0.55\left(\frac{\lambda\Phi_o'}{\Phi_o}\right)^{1/2} = 2{\cdot}10^{-4}\frac{\Phi_o'^{1/2}}{\Phi_o^{3/4}} \qquad [19{\cdot}98a]$$

$$d_{\min} = 2.2\left(\frac{\lambda\Phi_o}{\Phi_o'}\right)^{1/2} = 8{\cdot}10^{-4}\frac{\Phi_o^{1/4}}{\Phi_o'^{1/2}} \qquad [19{\cdot}98b]$$

To gain an idea of the magnitudes of these quantities, consider an electron microscope with an operating voltage $\Phi = 1000$ volts, a field of 1000 volts per centimeter at the surface of the specimen, and an initial velocity range $\Phi_o = 0.1$ volt, corresponding to thermionic emission. Then

$\alpha_o = 4{\cdot}10^{-2}$ radian $\qquad \alpha = 4{\cdot}10^{-4}$ radian $\qquad d_{\min} = 1.4{\cdot}10^{-5}$ cm

To attain this resolution the physical aperture of the objective, assuming that its focal length is 2 centimeters, would have to be about 0.015 millimeter ($2f\alpha$). It is clear that with these apertures the spherical aberration of the objective itself plays no role, so that its neglect in the derivation of the limit of resolution is justified.

It is also of interest to determine the limit of resolution without a limiting aperture in the objective. Then $\alpha_o = \pi/2$, $\alpha = (\Phi_o/\Phi)^{1/2}$, and, since $\Phi_r + \Phi_z = \Phi_o$, by Eq. 19·96

$$\Delta r = \frac{\Phi_o}{\Phi_o'} \qquad [19{\cdot}99]$$

By focusing the lens on the most suitable plane between the virtual object plane for rays with zero and that for rays with maximum initial axial velocity, this can be halved. Hence, the limit of resolution imposed by spherical and chromatic aberration becomes in this case

$$d_{\min} = \frac{\Phi_o}{\Phi_o'} + Cf\left(\frac{\Phi_o}{\Phi}\right)^{3/4} \qquad [19{\cdot}100]$$

Here also the second term, due to the spherical aberration of the lens, can normally be neglected. For the case considered before, $\Phi_o' = 1000$ volts per centimeter, $\Phi = 1000$ volts, $\Phi_o = 0.1$ volt,

$$d_{\min} = 10^{-4} + Cf{\cdot}10^{-6} \text{ cm}$$

The resolution is thus worse by a factor of about seven than that obtained with the lens provided with a limiting aperture of optimum size. With increasing field on the cathode, however, the resolution improves, and

the difference between the diaphragmed and undiaphragmed lens becomes rapidly smaller. A field of 100,000 volts per centimeter applied to the object is still quite reasonable. For this field Eq. 19·100 yields 100 A.U. for the diameter of the circle of confusion. The contribution of diffraction, which is $0.6\lambda = 24$ A.U. (sin $\alpha_o = 1$), is here of minor importance. This is not the case for extremely high fields, approaching the condition of cold emission. For such fields Recknagel[47] has derived wave-mechanically the relation

$$d_{\min} = \lambda^{5/4} \left(\frac{\Phi_o}{\Phi_o'}\right)^{1/4} \qquad [19\cdot101]$$

For a surface field corresponding to cold emission from tungsten ($3\cdot10^7$ volts per centimeter) this would yield as optimum resolution 12 A.U.

The limit of resolution for photoelectric and secondary-emission microscopes lies slightly higher than that of thermionic microscopes. This is a consequence of the greater initial velocities of the electrons. They necessitate the employment, for best results, of very narrow objective apertures, as is indicated by Eq. 19·98*a*.

REFERENCES

1. R. F. Baker, E. G. Ramberg, and J. Hillier, "The photographic action of electrons in the range between 40 and 212 kilovolts," *J. Applied Phys.*, Vol. 13, pp. 450–456, July 1942; Vol. 14, p. 39, January 1943.
2. N. F. Mott and H. S. W. Massey, *Theory of Atomic Collisions*, Oxford, 1933.
3. L. Marton and L. I. Schiff, "Determination of object thickness in electron microscope," *J. Applied Phys.*, Vol. 12, pp. 759–765, October 1941.
4. E. J. Williams, "Scattering of fast electrons and of cosmic rays particles," *Proc. Roy. Soc.*, London, Vol. A169, pp. 531–572, 1939.
5. S. Goudsmit and J. L. Saunderson, "Multiple scattering of electrons," *Phys. Rev.*, Vol. 57, pp. 24–29; Vol. 58, pp. 36–42, 1940.
6. H. Geiger and K. Scheel, *Handbuch der Physik*, Vol. 22, part 2, J. Springer, Berlin, 1933 (Second Edition).
7. G. Wentzel, "Dispersion of corpuscular rays as diffraction phenomenon," *Z. Physik*, Vol. 40, pp. 590–593, 1926.
8. H. Bethe, "Theory of the transmission of fast corpuscular rays through matter," *Ann. Physik*, Vol. 5, pp. 325–400, 1930.
9. R. Whiddington, "Transmission of cathode rays through matter," *Proc. Roy. Soc.*, London, Vol. A89, pp. 554–560, 1914.
10. P. Lenard, *Quantitatives über Kathodenstrahlen*, Heidelberg, 1918.
11. E. J. Williams, "The rate of loss of energy by β-particles in passing through matter," *Proc. Roy. Soc.*, London, Vol. A130, pp. 310–327, 1931.
12. P. White and G. Millington, "Velocity and distribution of β-particles after passing through thin foils," *Proc. Roy. Soc.*, London, Vol. A120, pp. 701–726, 1928.

[47] See reference 37.

13. H. Boochs, "Electron interferences," *Z. tech. Physik*, Vol. 19, p. 605, December 1938.
14. H. M. Terrill, "Absorption of cathode rays in aluminum," *Phys. Rev.*, Vol. 24, pp. 616–621, 1924.
15. O. Scherzer, "Theoretically attainable resolving power of the electron microscope," *Z. Physik*, Vol. 114, pp. 427–434, November 1939.
16. M. Born, *Optik*, Springer, Berlin, 1934.
17. J. Picht, "A general solution for the wave equation of an astigmatic beam," *Ann. Physik*, Vol. 77, pp. 685–782, October 1925.
18. R. Rebsch, "Theoretical resolving power of the electron microscope," *Ann. Physik*, Vol. 31, pp. 551–560, 1938.
19. E. Abbe, *Die Lehre von der Bildentstehung im Mikroskop*, edited by L. Lummer and F. Reiche, Braunschweig, 1910.
20. J. Hillier, "Fresnel diffraction of electrons as a contour phenomenon in electron supermicroscope images," *Phys. Rev.*, Vol. 58, p. 842, 1940.
21a. H. Boersch, "Fresnel diffraction phenomena in the supermicroscope," *Naturwissenschaften*, Vol. 28, p. 710, 1940.
21b. E. Ruska, "The origin of fringes about supermicroscopically imaged particles and their change with focusing," *Kolloid-Z.*, Vol. 105, pp. 43–52, 1943.
22. R. W. Wood, *Physical Optics*, The Macmillan Co., New York, 1934.
23. M. v. Ardenne, "The limits of the resolving power of the electron microscope," *Z. Physik*, Vol. 108, pp. 338–352, 1938.
24. M. v. Ardenne, "The determination of the resolving power of electron microscopes," *Physik. Z.*, Vol. 42, pp. 72–74, April 1941.
25. L. I. Schiff, "Ultimate resolving power of the electron microscope," *Phys. Rev.*, Vol. 61, pp. 721–722, 1942.
26. B. v. Borries and E. Ruska, "Effect of beam voltage on the supermicroscopic image," *Z. Physik*, Vol. 116, pp. 249–256, 1940.
27. J. Hillier and R. F. Baker, "The observation of crystalline reflections in electron-microscope images," *Phys. Rev.*, Vol. 61, pp. 722–723, June 1942.
28. R. D. Heidenreich, "Electron reflections in MgO crystals with the electron microscope," *Phys. Rev.*, Vol. 62, pp. 291–292, 1942.
29. M. v. Ardenne, "The electron scanning microscope — theoretical basis," *Z. Physik*, Vol. 109, pp. 553–572, 1938.
30. H. Boersch, "The electron shadow microscope I: geometric-optical experiments," *Z. tech. Physik*, Vol. 20, pp. 346–350, 1939.
31. R. A. Houstoun, *A Treatise on Light*, Longmans, Green and Co., London, 1924.
32. G. A. Morton and E. G. Ramberg, "The point projector microscope," *Phys. Rev.*, Vol. 56, p. 705, 1939.
33. E. Ruska, "The possibility of focusing cathode-ray pencils of large initial diameter," *Z. Physik*, Vol. 83, pp. 684–697, 1933.
34. G. Richter, "The velocity distribution of field electrons," *Z. Physik*, Vol. 119, pp. 406–414, 1942.
35. E. W. Müller, "Electron-microscope observation of field cathodes," *Z. Physik*, Vol. 106, pp. 541–550, 1937.
36. C. F. Eyring, S. S. Makeown, and R. A. Millikan, "Field currents from points," *Phys. Rev.*, Vol. 31, pp. 900–909, 1928.
37. A. Recknagel, "The resolving power of the electron microscope for emitters," *Z. Physik*, Vol. 120, pp. 331–362, 1943.

APPENDIX 1

THE PROBLEM OF NOISE IN AMPLIFICATION AND IN THE SCANNING MICROSCOPE

1. Reduction of Noise by the Employment of an Electron Multiplier as Preamplifier. Consider, in Fig. A1*a*, an ordinary phototube connected to a thermionic amplifier, the voltage across the coupling resistance R being impressed on the grid of the first tube of the amplifier.

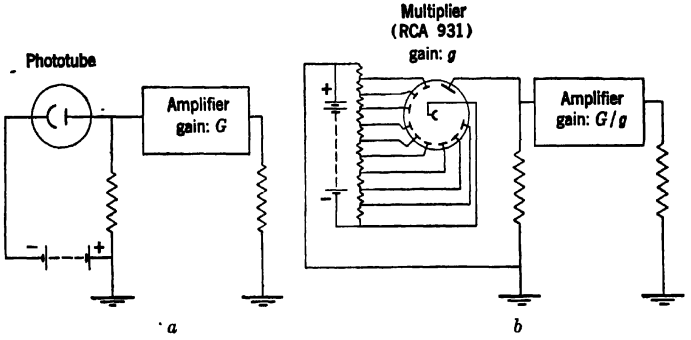


FIG. A1. The Amplification of Photocurrents with a Thermionic Amplifier (*a*) and an Electron Multiplier and Amplifier (*b*).

The total gain of the amplifier is assumed to be G and the range of frequencies which it is capable of transmitting, to lie between f and $f + \Delta f$. If the root-mean-square amplitude of the photocurrent, with a frequency lying in the above frequency band, is i_P, the signal-voltage output of the amplifier will be

$$V_s = G \cdot i_P \cdot R \qquad [A1]$$

At the same time the signal will be disturbed by random fluctuations in voltage, arising from two causes:

1. The fact that electron currents are made up of discrete, more or less random, pulses, corresponding to the emission of individual electrons. The theory of statistics teaches that the mean-square value of the resulting current fluctuations, taking account only of the fluctuation components in the transmitted frequency band of width Δf, is

$$\overline{i_n^2} = 2ei_P\Delta f = 32 \cdot 10^{-20} i_P \Delta f \text{ amp}^2 \qquad [A2]$$

2. The existence of voltage fluctuations across the terminals of a resistor, resulting from the thermal motions of the electrons and ions within. Their mean-square value at room temperature ($T = 300$ degrees absolute) is given by

$$\overline{V_n^2} = 4kTR\Delta f = 1.6\cdot10^{-20}R\Delta f \text{ volt}^2 \qquad [A3]$$

Thus the output noise will be given by

$$\overline{V_n^2} = G^2(32\cdot10^{-20}i_P\Delta fR^2 + 1.6\cdot10^{-20}R\Delta f) \qquad [A4]$$

The purity and recognizability of the signal depend on the ratio S of the signal to the noise voltage:

$$S = \frac{V_s}{(\overline{V_n^2})^{1/2}} = \frac{i_P}{\{(32\cdot10^{-20}i_P + 1.6\cdot10^{-20}R^{-1})\Delta f\}^{1/2}} \qquad [A5]$$

The resistance R has an effective upper limit, set by the top frequency of the transmitted band and the capacities of the phototube collector and the input grid of the first tube of the amplifier, these determining the magnitude of the capacitative shunt impedance of the resistor. Assume, somewhat arbitrarily, that $R = 10{,}000$ ohms and that $\Delta f = 1000$ cycles per second. Then, if $S = 10$, $i_P \cong 4\cdot10^{-10}$ ampere. The term due to shot noise is quite negligible, being smaller by a factor $20i_PR \cong 10^{-4}$ than the term due to resistor noise. The resistor noise (together, eventually, with the noise introduced by the input tube) is quite generally the dominant factor in the noise output of a conventional amplifier.

Next, consider, in Fig. A1*b*, a multiplier of gain g, with a similar photocathode, coupled to an amplifier of gain G/g with the same transmitted frequency band as before by means of a similar resistance R. The signal output will again be given by Eq. A1. If the gain per stage of the multiplier is great compared to unity, the shot noise in the original current emitted by the photocathode, $i_P + i_d$,[1] is simply multiplied along with the current itself.[2] Thus now

$$\overline{V_n^2} = \left(\frac{G}{g}\right)^2 (32\cdot10^{-20}(i_P + i_d)g^2\Delta fR^2 + 1.6\cdot10^{-20}R\Delta f)$$
$$= G^2\left(32\cdot10^{-20}(i_P + i_d)\Delta fR^2 + 1.6\cdot10^{-20}R\frac{\Delta f}{g^2}\right) \qquad [A6]$$

[1] i_d here denotes the dark current referred back to the cathode and is of the order of $10^{-6}/g = 5\cdot10^{-12}$ amp for the RCA 931 operated at 1250 volts (Janes and Glover, reference 1).

[2] See Zworykin, Morton, and Malter, reference 2.

Apart from the replacement of i_P with $i_P + i_d$, the effect of inserting the multiplier in the place of the phototube is to reduce the thermal noise by a factor g^2, which is usually of the order of 10^{10} to 10^{12}. Hence the thermal noise is now rendered insignificant, and the signal-to-noise ratio becomes

$$S = \frac{i_P}{[32 \cdot 10^{-20}(i_P + i_d)\Delta f]^{1/2}} \qquad [A7]$$

With $i_d = 5 \cdot 10^{-12}$ ampere and $S = 10$, $i_P = 4 \cdot 10^{-13}$ ampere. Thus the multiplier decreases, under the circumstances here considered, the least signal which is transmitted satisfactorily by three orders of magnitude. Equation A6 also shows that for small values of g, such as are obtained with a single-stage multiplier (for instance, the orbital beam multiplier), the gain in signal-to-noise ratio is approximately equal to g.

2. Noise in the Scanning Microscope. Consider, first, a scanning microscope for continuous observation such as that sketched in Fig. 3·13. Assume, furthermore, that the signal is amplified by a conventional resistance-coupled amplifier. Then, by Eq. A5, the signal-to-noise ratio is given by

$$S = \frac{i_P R^{1/2}}{(1.6 \cdot 10^{-20}\Delta f)^{1/2}} \qquad [A8]$$

The shot noise is negligible compared with the thermal noise introduced by the coupling resistor R. The band width Δf transmitted by the amplifier must be such that variations in intensity from one picture element to the next are transmitted to the viewing tube. This signifies a top frequency of the order of the reciprocal time of scanning for a single element. For a standard 441-line television image, the number of elements is (for a width-to-height ratio 4:3) $441 \cdot 441 \cdot 4/3 = 260{,}000$. Since the whole picture is scanned 30 times a second, the required top frequency would appear to be 7.8 megacycles per second. A more detailed examination of the problem indicates, however, that a band width of 4 megacycles per second — assigning a half period of the top frequency to the scanning of a picture element — is sufficient.[3]

Furthermore, in the actual coupling circuits employed (for example, in television practice), the capacitative conductance of the collector and input-grid capacitances greatly exceeds, at the higher frequencies, the parallel conductance of the coupling resistance. The resulting distortion of the frequency response of the coupling circuit is corrected by a compensating network in a higher stage of the amplifier.

Under these most favorable circumstances, the signal-to-noise ratio is

[3] See Zworykin and Morton, reference 3, section 6·8.

determined by the shunt capacitance and the equivalent noise resistance of the first tube of the amplifier rather than by the coupling resistance. A calculation[4] shows that, if the sum of the collector and input-grid capacitances is 15 micromicrofarads and the equivalent noise resistance of the first tube is 1000 ohms, the signal-to-noise ratio obtained is the same as for a coupling resistance $R = 7000$ ohms without shunt capacitance and with a noise-free amplifier (Eq. A8). Accordingly,

$$S = 3.3 \cdot 10^8 i_P \qquad [A9]$$

i_P is here the signal current received by the collector. If the secondary-emission ratio of the object approximates unity, as is commonly the case, the signal current i_P may be set equal to the beam current i_b. The beam current, by the law of Helmholtz-Lagrange (Eq. 10·32), has an upper limit

$$i_b = j_b \cdot A = \frac{\pi d^2}{4} \frac{11{,}607V}{T} \alpha^2 j_o \qquad [A10]$$

Here A is the cross-section area of the probe, d its diameter, T the temperature of the cathode in degrees absolute, eV the kinetic energy of the probe electrons, α the angle of convergence of the electrons in the probe, and j_o the emission per unit area of the source.

For resolving powers much inferior to the optimum attainable with the lenses employed, the maximum value of d is fixed by the spherical aberration of the final lens, which may be written (see section 17·2):

$$d = Cf\alpha^3 \qquad [A11]$$

so that the relation between the beam current and the attainable resolving power becomes

$$i_b = \frac{\pi}{4} \cdot \frac{d^{8/3}}{(Cf)^{2/3}} \cdot \frac{11{,}607V}{T} \cdot j_o \qquad [A12]$$

By substituting $i_b = i_P = 3 \cdot 10^{-9} S$ from Eq. A9 and solving for d with $Cf = 10$ centimeters,[5] $T = 1000°$ A, $j_o = 1$ ampere per square centimeter (these values applying approximately for oxide-coated cathodes), $V = 800$ volts, and $S = 10$,

$$d = 10^{-4} \text{ cm} = 1 \text{ micron}$$

The proper value of the factor S depends on the variation of the secondary-emission factor over the object structure examined. The signal-to-noise ratio must be large enough for the variation in signal

[4] See Zworykin and Morton, reference 3, section 14·17.

[5] The value $Cf = 10$ cm assumes electrostatic lenses. A lower value, and hence a somewhat more favorable signal-to-noise ratio, is attainable with magnetic lenses.

strength at different parts of the structure to overbalance by a considerable factor — at least 3 — the random signal fluctuations. For example, if the variation in secondary emission amounts to only 1 per cent, S should have a value at least equal to 300, and the optimum resolution is given by

$$d = 1 \text{ micron} \cdot (30)^{\frac{2}{3}} = 4 \text{ microns}$$

In von Ardenne's transmission scanning microscope (section 3·4), which does not possess an amplifier, so that only shot noise affects the image, noise does not limit the resolution of the instrument. If it is assumed, for the magnetic objective, that $Cf = 1$ centimeter and $V = 23$ kilovolts, Eq. 19·50 yields for the optimum resolution and the corresponding angular divergence

$$d_{\min} = 15 \cdot 10^{-8} \text{ cm} \qquad \alpha = 4.5 \cdot 10^{-3} \text{ radian}$$

so that the beam current, as given by Eq. A10, becomes $i_b = 9.6 \cdot 10^{-14}$ ampere. For a square 300-line picture recorded in 10 minutes, the time per picture element is $t_e = 10 \cdot 60/(300)^2 = 6.7 \cdot 10^{-3}$ second, and the charge per element transmitted through a clear part of the object

$$i_b \cdot t_e = 6.4 \cdot 10^{-16} \text{ coulomb} = 4000 \text{ electrons}$$

Since the mean-square fluctuation in the number of particles emitted at random is equal to this number, the signal-to-noise ratio becomes

$$S = \frac{n}{n^{\frac{1}{2}}} = 63$$

Consider next the scanning microscope for surface observation employing multiplier preamplification and facsimile recording.[6] Here two new sources of noise must be considered. Thus, if the secondary-emission ratio of the object approximates unity, the mean-square noise in the beam of secondary electrons will be about twice as great as that in the primary beam. A further doubling of the beam noise must be expected — if the photocurrent in the multiplier is about equal to the secondary-electron current incident on the fluorescent screen — as the result of the translation of the secondary current into light and of the light, in turn, into multiplier photocurrent. However, the thermal noise of the coupling resistor is rendered innocuous by the reduction in the amplification of the conventional amplifier. Accordingly, except for a factor 4 in the denominator, the signal-to-noise ratio is given by Eq. A7:

$$S = \frac{i_s}{[32 \cdot 10^{-20}(4i_s + i_d)\Delta f]^{\frac{1}{2}}} \qquad [A13]$$

[6] See Zworykin, Hillier, and Snyder, reference 4.

Since the facsimile receiver records a 750-line image ($7\frac{1}{2} \times 6$ inches^2) in 10 minutes, the transmitted band width Δf must be, if no resolution is to be lost in the transmission process (in analogy to the system for continuous observation previously discussed),

$$\frac{1}{2}\,\frac{750 \cdot 750}{60 \cdot 10}\,\frac{7.5}{6} = 580 \text{ cycles per sec}$$

Actually, a transmission band of 700 cycles per second is employed. For the specially selected multiplier used, i_d, that is, the current leaving the photocathode which, after multiplication, would give rise to the observed dark current in the output of the multiplier, was only $5 \cdot 10^{-13}$ ampere. Thus

$$S = \frac{i_s}{[2.24 \cdot 10^{-16}(4i_s + 5 \cdot 10^{-13})]^{\frac{1}{2}}}$$

For a velocity of the probe electrons corresponding to 800 volts, with $Cf = 10$ centimeters, by Eq. 19·50

$$d_{\min} = 91 \cdot 10^{-8} \text{ cm} \qquad \alpha = 3.8 \cdot 10^{-3} \text{ radian}$$

Hence, by Eq. A10, $i_s = i_b = 8.7 \cdot 10^{-14}$. This leads to a value for the signal-to-noise ratio under ideal conditions of

$$S = 6$$

If account is taken of the larger thermionic emission of the tungsten cathode (as compared with $j_o = 1$ ampere per square centimeter) the figure becomes more favorable, even though T is larger than assumed. The system should thus be capable of reproducing images with the maximum resolution of which the electron lenses forming the probe are capable.

REFERENCES

1. R. B. Janes and A. M. Glover, "Recent developments in phototubes," *RCA Review*, Vol. 6, pp. 43–54, July 1941.
2. V. K. Zworykin, G. A. Morton, and L. Malter, "The secondary-emission multiplier — a new electronic device," *Proc. Inst. Radio Engrs.*, Vol. 24, pp. 351–375, March 1936.
3. V. K. Zworykin and G. A. Morton, *Television*, John Wiley and Sons, New York, 1940.
4. V. K. Zworykin, J. Hillier, and R. L. Snyder, "A scanning electron microscope," *ASTM Bull.*, No. 117, pp. 15–23, August 1942.

APPENDIX II

MISCELLANEOUS TABLES

TABLE I

Fundamental Physical Constants

(From R. T. Birge, "New Table of Values of the General Physical Constants [as of August 1941]," *Reviews of Modern Physics*, Vol. 13, pp. 233–239, October 1941.)

velocity of light, c	$2.99776 \pm 0.00004 \cdot 10^{10}$ cm·sec^{-1}
charge of the electron, e	$4.8025 \pm 0.0010 \cdot 10^{-10}$ e.s.u.
mass of the electron, m	$9.1066 \pm 0.0032 \cdot 10^{-28}$ g
specific charge of the electron, e/m	$5.2737 \pm 0.0015 \cdot 10^{17}$ e.s.u.
mass of the proton, M_p	$1.67248 \pm 0.00002 \cdot 10^{-24}$ g
ratio of mass of proton to mass of electron, M_p/m	1836.56 ± 0.56
atomic weight of proton (on chemical scale)	1.0073 ± 0.0003
Planck constant, h	$6.624 \pm 0.002 \cdot 10^{-27}$ erg·sec
Boltzmann constant, k	$1.38047 \pm 0.00026 \cdot 10^{-16}$ erg·deg^{-1}
Avogadro number, N_o	$6.0228 \pm 0.0011 \cdot 10^{23}$ mole^{-1}
Loschmidt number, n_o (for perfect gas at 0° C and atmospheric pressure)	$2.6870 \pm 0.0005 \cdot 10^{19}$ cm^{-3}

TABLE II

Some Useful Derived Constants

velocity of a 1-volt electron, $(2e \cdot 10^8/(mc))^{1/2}$	$5.932 \cdot 10^7$ cm·sec^{-1}
de Broglie wave length of a 1-volt electron, $hc^{1/2}/(2em \cdot 10^8)^{1/2}$	$1.2262 \cdot 10^{-7}$ cm
temperature of thermionic emitter for which mean energy of emission normal to the surface is 1 electron-volt, $e \cdot 10^8/(kc)$	11,607° A
radius of curvature of path of 1-volt electron in a magnetic field of 1 gauss normal to it, $(2mc \cdot 10^8/e)^{1/2}$	3.3715 cm
frequency of rotation of electron in a magnetic field of 1 gauss, $e/(2\pi mc)$	$2.8003 \cdot 10^6$ sec^{-1}

TABLE III

Conversion Factors

A. Units of Length:

1 cm = 10 mm = 10^4 microns (μ) = 10^7 millimicrons (mμ)
= 10^8 Angstrom Units (A.U.)
1 ft = 12 in. = 30.4801 cm
1 mil = 0.001 in. = 25.4 microns

B. Electrical Units:

1 coulomb = $c/10 \sim 3 \cdot 10^9$ e.s.u. of charge
1 ampere = $c/10 \sim 3 \cdot 10^9$ e.s.u. of current
1 volt = $10^8/c \sim 1/300$ e.s.u. of electric potential
1 ohm = $10^9/c^2 \sim 10^{-11}/9$ e.s.u. of electric resistance
1 gauss = 1 e.m.u. of magnetic field

TABLE IV

MEAN FREE PATH OF ELECTRONS IN AIR AS FUNCTION OF PRESSURE

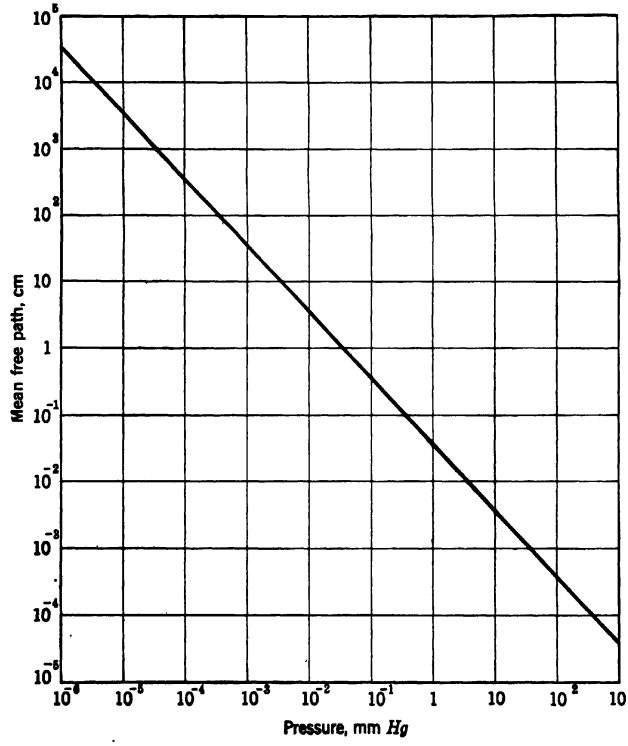


TABLE V

DE BROGLIE WAVE LENGTHS OF VARIOUS PARTICLES IN MOTION

Particle	Mass, g	Velocity, cm/sec	Wave Length, A.U.
1,000,000-volt electron	$2.694 \cdot 10^{-27}$	$2.821 \cdot 10^{10}$	0.008715
300,000-volt electron	$1.446 \cdot 10^{-27}$	$2.328 \cdot 10^{10}$	0.01968
50,000-volt electron	$1.000 \cdot 10^{-27}$	$1.237 \cdot 10^{10}$	0.05354
1000-volt electron	$9.125 \cdot 10^{-28}$	$1.873 \cdot 10^{9}$	0.3876
1-volt electron	$9.107 \cdot 10^{-28}$	$5.932 \cdot 10^{7}$	12.262
1,000,000-volt proton	$1.674 \cdot 10^{-24}$	$1.382 \cdot 10^{9}$	0.0002860
50,000-volt proton	$1.672 \cdot 10^{-24}$	$3.095 \cdot 10^{8}$	0.001280
1-volt proton	$1.672 \cdot 10^{-24}$	$1.384 \cdot 10^{6}$	0.2861
Baseball	140	2000	$2.365714 \cdot 10^{-24}$

AUTHOR INDEX

Abbe, E., 81, 703
Adams, E. P., 369, 401, 416, 429, 521, 526, 533, 547, 567, 570, 581, 583, 586, 652
Allison, S. K., 112
Alvarez, L. W., 74
Anderson, T. F., 283, 286, 287, 292, 294, 296, 298, 301, 303, 304, 305 310
von Ardenne, M., 92, 100, 101, 141, 144, 160, 161, 167, 200, 241, 245, 246, 257, 261, 266, 290, 308, 309, 310, 320, 325, 621, 724, 727, 736
Aston, F. W., 69, 71
Attwood, S. S., 480

Bachman, C. H., 95
Bainbridge, K. T., 71
Baker, R. F., 262, 273, 328, 674, 727, 734
Banca, M. C., 173
Barnes, R. B., 259, 315, 326
Beard, D. 290
Beard, J. W., 290
Becker, H., 621
Behne, R., 117
Beischer, D., 325
Bender, J. F., 173
Benjamin, M., 120
Bennek, H., 248
Bertram, S., 381, 450
Bethe, H. A., 667, 686
Binks, W., 222
Birge, R. T., 753
Bode, H., 220
de Boer, J. H., 43
Boersch, H., 110, 111, 716, 736
Boochs, H., 690
Born, M., 693
von Borries, B., 163, 176, 180, 189, 194, 205, 258, 261, 269 270, 272, 311, 312, 337, 435, 591, 594, 733
Bothe, W., 260, 682, 685, 688
Bowman-Manifold, M., 392
Brobeck, W. M., 74
de Broglie, L., 84, 345
Brüche, E., 93, 113, 114, 116, 376, 383, 435
Bublitz, G., 488
Burgers, W. G., 114
Burton, C. J., 259, 315, 326
Burton, E. F., 173
Busch, H., 493

Calbick, C. J., 445
Caldwell, O. G., 320
Chambers, L. A., 292
Clark, G. L., 220, 222
Coates, W. M., 73
Coggeshall, N. D., 71
Compton, A. H., 112
Cooksey, D., 73, 74
Corson, D. R., 74
Coslett, V. E., 69
Craig, P. H., 489

Dahl, O., 71
Davisson, C. J., 445
Delbrück, M., 287
Dempster, A. J., 69, 71
Dosse, J., 140, 205, 489, 594, 616, 621, 627, 643
Driest, E., 173
Duffendack, O. S., 328

Eitel, W., 268, 321
Ellis, C. D., 67
Emde, F., 376, 474, 536
Epstein, D. W., 39, 451, 518, 617, 621
Eyring, C. F., 108, 742

Farnsworth, P. T., 49
Ferris, W. R., 63
Field, L. M., 618, 621
Frevel, L. K., 271
Friedrich-Freksa, H., 241, 245, 309
Fullam, E. F., 308

Gabor, D., 417
Gans, R., 412, 414, 445
Geiger, H., 682, 688
Gessler, A. E., 308
Glaser, W., 474, 506, 543, 559, 572, 605, 606, 616, 617, 624, 629, 630
Glöde, H., 220
Glover, A. M., 62, 748
Gölz, E., 91, 93, 331
Gotthardt, E., 268, 321
Goudsmit, S., 675, 685
Green, R. H., 286
Gundert, E., 395, 617, 621

Haefer, R., 108, 120, 330
Haeff, A. V., 594
Hafstad, L. R., 71
Hall, C. E., 173, 305, 322
Hamilton, W. R., 6
Hamly, D. H., 308
Hammond, W. M., 309
Hanawalt, J. D., 271
Hance, R. T., 305
Hansen, W. W., 655
Harker, D., 244, 246, 249
Hass, G., 245
Heidenreich, R. D., 251, 331, 334, 335, 735
Heinze, W., 114
Heisenberg, W., 349
Henneberg, W., 267
Herzog, R., 70
Hickok, W. H., 41
Hillier, J., 102, 159, 165, 170, 173, 177, 201, 233, 234, 262, 273, 328, 674, 716, 734, 751
Hipple, J. A., 71
Holst, G., 43
Houstoun, R. A., 345, 737
Houtermans, F. G., 113
Hughes, A. L., 68
Hull, A. W., 528
Humbert, R. P., 320
Husemann, E., 310

Iams, H., 41, 50, 521

Jahnke, E., 376, 474, 536
Jakob, A., 294
Jakus, M. A., 305
Janes, R. B., 41, 62, 748
Janzen, S., 337
Jeans, J., 368, 468
Jenkins, R. O., 120
Johannson, H., 113, 114, 117
Johnson, J. B., 591
Johnson, R. P., 115
Jordan, E. B., 71
Jung, F., 300

Kausche, G. A., 283, 286, 288, 311, 312
Kehler, H., 245
Kerst, D. W., 76, 668
Kinder, E., 175
Kirchner, F., 242
Klemperer, O., 68
Kleynen, P. H. J. A., 421
Knecht, W., 113, 114
Knoll, M., 57, 113, 116, 433, 441
Koch, H. W., 313
Kormes, M., 387
Kregel, E. A., 326
Kretschmer, M., 325
Krumm, J., 176

Lackman, D. B., 291
Ladd, W. A., 313
Langmuir, D. B., 356, 417
Law, R. R., 38, 41
Lawrence, E. O., 72, 73, 74, 660
Lembke, A., 294
Lenard, P., 688, 690
Livingston, M. S., 73
Luria, S. E., 286

Mahl, H., 89, 91, 92, 94, 95, 98, 110, 114, 116, 154, 219, 247, 248, 267, 294, 315, 326, 328, 335
Makeown, S. S., 108, 742
Maloff, I. G., 518
Malsch, F., 126
Malter, L., 61, 520, 748
Marshall, C. E., 320
Martin, L. C., 173
Martin, S. T., 115, 120
Marton, L., 158, 173, 267, 282, 324, 675, 682
Massey, H. S. W., 675, 704
Mattauch, J., 70
McBain, J. W., 324

McKinley, G. M., 308
McMillen, J. H., 114
Mecklenburg, W., 113
Mehl, R. F., 331
Meschter, E., 117
Meyer, C. F., 276
Miller, D. C., 345
Millikan, R. A., 108, 742
Millington, G., 688
Morton, G. A., 43, 50, 61, 109, 520, 579, 585, 637, 738, 748, 749, 750
Morton, H. E., 293, 294, 296
Mott, N. F., 675, 704
Mudd, S., 291, 292, 296, 298
Müller, E. W., 108, 119, 120, 741
Müller, H. O., 173, 176, 201, 269, 270, 308, 321
Mutscheller, A., 220

Nelson, H. R., 249
Nicoll, F. H., 392, 461

O'Brien, H. C., 308
O'Daniel, H., 272, 321
Ollendorf, F., 368
Osgood, W. F., 527

Paehr, W., 130
Painter, W. H., 37
Parnum, D. H., 173
Pasewaldt, C. W. A., 308
Peck, V. G., 251, 331, 334
Pendzich, A., 175
Pfankuch, E., 283, 285
Picard, R. G., 328
Picht, J., 696
Pierce, J. R., 647
Plass, G. N., 450, 609, 611
Ploos van Amstel, J. J. A., 114
Pohl, J., 116
Polevitsky, K., 292
Prebus, A., 173, 272
Pyl, G., 290

Radczewski, O. E., 272, 321
Rajchman, J., 62, 415, 421
Ramberg, E. G., 109, 248, 262, 331, 575, 579, 585, 609, 637, 674, 738
Ramo, S., 95
Rebsch, R., 604, 608, 617, 628, 698, 726
Recknagel, A., 414, 461, 569, 605, 745
Richards, A. G., 301, 303, 304, 305
Richter, G., 740
Riedl, H., 634
Rinn, H. W., 271
Rittenberg, D., 71
Robinson, H., 67
Rochow, T. G., 259
Rogowski, W., 543
Rojansky, V., 68
Rose, A., 47, 521, 527, 531
Rose, M. E., 665, 667, 668
Rüdiger, O., 248
Ruedy, J. E., 328
Ruska, E., 116, 134, 136, 137, 139, 163, 172, 173, 176, 180, 189, 194, 201, 205, 224, 258, 269, 270, 272, 433, 435, 591, 621, 716, 733, 739
Ruska, H., 283, 288, 294, 301, 310, 325

Salisbury, W. W., 74
Saunderson, J. L., 675, 685
Schade, O. H., 57
Schaefer, V. J., 244, 246, 249, 251
Scheel, K., 682, 688
Scherzer, O., 376, 383, 402, 543, 562, 571, 604, 608, 634, 691, 736, 738
Schiff, L. I., 159, 267, 675, 682, 729
Schmieder, F., 316
Schmitt, F. O., 305
Schneider, W., 608, 617
Schoen, A. L., 173, 322
Schoon, Th., 313
Schroeder, A. C., 144
Schulze, W., 113
Schusterius, C., 321
Schwartz, E., 130
Schwarzschild, K., 353
Scott, G. H., 114
Sears, G. R., 326
Serber, R., 668
Sharp, D. G., 290
Shaw, B. T., 320
Shortley, G. H., 388
Simard, G. L., 249
Sloan, D. H., 72, 73, 660
Smadel, J. E., 286
Smith, G. S., 488
Snyder, R. L., 102, 751
Spangenberg, K., 618, 621

Stäblein, F., 248
Stanley, W. M., 283, 287, 310
Starling, S. G., 67
Steinbach, H. B., 305
Stranski, I. N., 335
Strehl, K., 555
Strübig, H., 130

Taylor, A. R., 290
Terrill, H. M., 690
Teves, M. C., 43
Thomas, L. H., 668
Thompson, H. C., 57
Thomson, J. J., 8, 66, 344, 520
Thornton, R. L., 74
Tuve, M. A., 71

van Urk, A. Th., 485
Vance, A. W., 165, 201, 225, 233, 234
Veenemans, C. F., 43
Voit, H., 634, 635
Vold, R. D., 324
Volk, K. E., 248

Wagener, S., 114
Wagner, H. M., 63
Wallraff, A., 621
Watson, E. E., 594
Watson, J. H. L., 308
Webster, D. L., 655
Weichardt, H., 441
Weller, R., 388
Wentzel, G., 683
Whelpton, R. V., 173
Whiddington, R., 687
White, P., 688
Wideroe, R., 76
Wiechert, E., 493
Wiegand, W. R., 313
Williams, E. J., 675, 685, 688
Wilson, R. R., 667
Wolpers, C., 301
Wood, R. W., 717
Woods, F. S., 400

Young, C. J., 102
Young, L. B., 71

Zworykin, V. K., 33, 35, 37, 40, 43, 50, 61, 62, 102, 117, 165, 170, 177, 201, 233, 248, 273, 328, 331, 421, 520, 579, 748, 749, 750, 751

Aberrations, astigmatism, 23, 209, **551**
correction, 635
chromatic, 25, 124, 157, 206, 210, **570**, 720
correction, 641
image contours, 710
coma, 22, 210, **553**
correction, 633
curvature of field, 22, 209, **550**
correction, 635
distortion, 24, 209, **549**
correction, 639
relativistic, 215, 653
space charge, 27, 591, 647
spherical, 22, 123, 205, **554**, **603**, 696
contrast, 704
correction, 604, 630
determination, 617
Fresnel diffraction, 719
image contours, 708
short lens, 616
voltage dependence, 650
weak lens, 607
total, 549, 725
Aberrations, classification, 22, 540
Aberrations, geometrical, 22, 541
evaluation of constants, 555
Aberrations, lenses, bell-shaped field, 624
cathode lens, 577
combined uniform fields, 538, 623, 644
electron mirror, 564, 575
high voltages, at, 657
image tube, 637, 645
objective lens, 146
projector lens, 147
uniform electric field, 623
uniform magnetic field, 499, 623, 644
Absorption of electrons, 7, 674, 690
Accelerating devices, 71, 660
Accelerating potential, *see* Power supply
Adjustment, electron microscope, alignment, 252, 600
gun, 130
Air lock, 93, 186; *see also* Specimen chamber and Plate chamber
Airy disk, 81, 694
Alignment, electron microscope, 252, 600
Ampère's law, 466, 470
Ampere turns, 135
Analogue, electron and light microscopes, 86, 672
light and electron optics, 4, 350
Antibodies, 288
Aperture, angular, 82, 205
control, 256
condenser, 182
objective, 182, 674
large, 701
small, 691
projector, 148
shielding, 181
Aperture defect, *see* Aberrations, spherical
Aperture lenses, 434, 441
Application of electron microscope, 281
Astigmatism, *see* Aberrations
Atoms, observation of, 159, 729
Axially symmetric systems, lens properties, 423, 499; *see also* Electrostatic lens, Magnetic lens, and Trajectories
potential distribution (electrostatic), 375

Bacteria with electron microscope, 282, 290
living, 241, 309
reaction with antiserum, 298
Bacteriophage, 285
Beam, electron, cathode ray tube, 31
diameter, 39, 41
electron microscope gun, 128
intensity, 354
space-charge limited, 647
Beam tubes, 55
Bell-shaped magnetic field, 506, 624
Betatron, 76, 668

Biological research, electron microscope, 282
Biot-Savart's law, 468
Blood, 300
Bragg's law, 733

Carbon black, 312
Cardinal points, electron lens, 426
Cathode, 4, 34; *see also* Emission and Gun
  current supply, 238
  electron microscope, 127
    adjustment, 254
  life, 127
  ray paths, 403, 415
Cathode lenses, 18, 434, **452, 577**
Cellulose, 325
Cement, 321
Central pencils, 88
Chemical research, electron microscope, 310
Chromatic aberration, *see* Aberrations
Circle of confusion, *see* Aberrations
Clays, 320
Coaxial cylinders, potential distribution, 378, 394
  ray paths in, 411
Coil design, 144
Cold emission, *see* Field emission
Collagen, 305
Collodion films, preparation, 242
Colloidal particles, 246, 311
Columbian Carbon Company, 246, 258, 313
Coma, *see* Aberrations
Condenser lens, electron microscope, 42, 87, **131**, 181
    aperture, 132
    pole pieces, 134
Conformal mapping, 360
Contamination, effect on electron microscope, 160
Contour effects, 708
Contrast, electron microscope image, focusing, 255
    illuminating aperture, 714
    large aperture, 710
    small aperture, 700
    specimen thickness, 158
    spherical aberration, 704
Contrast perception, 158
Convergence, ray, 402, 582
Crossed electric and magnetic fields, 520
Crystalline diffraction in electron microscope image, 733
Current regulation, lens coil, 124, 237
Curvature of field, *see* Aberrations
Cyclotron, 73, 664
Cylindrical lens (two-dimensional), electrostatic, field distribution, 372
  magnetic, 518

Dark-field illumination, 133, 731
Deflection of electron beam, in crossed fields, 523
  in electric field, 52
  in magnetic field, 53, 516
Depth of field, 155
Diatoms, 308
Differences, method of, 407
Diffraction camera, adaptation of electron microscope, 271
Diffraction of electrons, electron microscope image, crystals, 733
  Fresnel, electron microscope, 716
  lens aperture, 82, 205, 693
Diffraction of light, 5, 345, 703
Diffusion pump, 196
Dissector tube, 49
Distortion, *see* Aberrations
Disturbing fields, 125, **159**, 217, 238, **597**

Earth's field, effect on electron microscope, 161
Elastic scattering, 675
Electrolytic plotting, electric fields, 389, 647
  magnetic fields, 477
Electron, 1, 344
  absorption, 7, 674, 690
  analogy with light, 4, 344, 350
  diffraction, *see* Diffraction
  emission, *see* Emission
  energy losses, 676, 686
  index of refraction, 7, 30, 352
  motion, *see* Trajectories
  penetration, 7, 688
  scattering, *see* Scattering
  velocity, 30, 754
    initial, 41, 126, 207

Electron, velocity, specimen thickness, effect of, 208
wave aspect, 344
wave length, 5, 84, **350**, 754
Electron gun, *see* Gun
Electron lens, *see also* Condenser lens, Electrostatic lens, Magnetic lens, Objecti · lens, Projector lens, and Screen lenses
aberrations, *see* Aberrations
cardinal points, 426
focal length, 14, 140, 151, 427
magnetic and electric combined, **499**, **535**, 540, 656
magnification, *see* Magnification
merit, figure of, 608
properties, 18, 21, **423**, **499**
refractive power, 136
Electron microscope, adaptation for diffraction, 270
application of, 281
comparison, magnetic and electrostatic, 98
general discussion, 81
image formation, 87, **672**, 691, 701
magnification, *see* Magnification
operation, 199; *see also* Adjustment, Focusing, Photographic plate, Specimen, Specimen preparation
resolving power, *see* Resolving power
tolerances, construction and alignment, 216
vacuum system, 165, 172, 195
voltage regulation, electrostatic, 90, 215
magnetic, 124, 213
Electron microscope, electrostatic, 89
A.E.G., 91
General Electric, 95
Electron microscope, emission, 81, 112, 741
field emission, 119, 741
photoelectric emission, 116, 742
projection, 115
secondary emission, 117, 742
thermionic emission, 113, 742
Electron microscope, magnetic, von Ardenne universal, **167**, 179, 183, 188, 193, 198
Electron microscope, magnetic, high voltage, 200
RCA Small, **170**, 181, 184, 189, 194, 199
power supply, 236
RCA Type A, 173
RCA Type B, **163**, 177, 181, 186, 191, 195, 199
power supplies, 234, 237
Ruska, early magnetic, 172
Siemens Supermicroscope, **163**, 179, 183, 193, 198
Toronto, 173
yoke type, 175
Electron microscope, point projector, 106, 738
Electron microscope, scanning, 98, 735
Electron microscope, shadow, 106, 736
Electron mirror, 19, 456
chromatic aberration, 575
correction of aberration with, 29, 630, 645
geometrical aberrations, 564
refractive power, 463
Electron multiplier, *see* Secondary emission multiplier
Electron optics, 343, 351, 353
analogy with light optics, 4, 86, **344**, 350
Electron probe, 99, 736
Electrostatic fields, *see* Potential distribution
Electrostatic lens, aberrations, *see* Aberrations
cathode, 18, 434, **452**, **577**
classification, 433
Gaussian dioptrics, 423
immersion, 434, 447
aberrations, 608; *see also* Aberrations
negative, 433, 441
ray equation, 401, 429
relativity correction, 651
slit lens, 368, 372
unipotential, 20, 30, 90, 434, **437**
aberrations, 608; *see also* Aberrations
Emission of electrons, cold, 4
field, 4
gas discharge, 4, 126

Emission of electrons, photoelectric, 4
secondary, 4, 59, 99
specific, 127, 354
thermionic, 3, 127, 207
Equipotential, *see also* Potential distribution
definition, 11
intersection, 373, 377
magnetic, 479, 512
Evaporated films, 328
Exit pupil, 424

Fermat's law, 351
Ferromagnetic material, 475
Fibrin, 301
Field aberrations, *see* Aberrations
Field distribution, magnetic lens, 139, 152, 490
Field emission, 4
Field emission microscope, 119, 741
Filament, *see* Cathode
Fluorescent screen, 95, 168, 171, 191
single crystal, 257
Focal length, 427; *see also* corresponding lenses
definition, 14
measurement, 140, 151
Focal points, 142, 426
definition, 12
multiple, 142
Focusing electron microscopes, electrostatic, 93
magnetic, 255
scanning microscope, 104
Form factor, 135
Formvar, 244
Fraunhofer diffraction, 693
Fresnel diffraction, 716, 737, 739, 741

Gans, path determination, 412
Gas discharge, 4, 126
Gaussian dioptrics, electrostatic lens, 423
Geometric aberration, *see* Aberrations
Graphical methods, ray path determination, combined magnetic and electrostatic fields, 512, 535
electrostatic field, 414
magnetic field, 512
Graphical potential mapping, Liebmann procedure, 386
Grid, focusing, 128
Gun, electron, 33; *see also* Cathode
electron microscope, 42, 125, 177
adjustment, 252
high voltage, 201
focusing grid, 128

Hamiltonian Characteristic function, 543
Helmholtz-Lagrange principle, 355
Huygens' principle, 345

Iconoscope, 33, 50
Illumination, electron microscope, angular aperture, 131, 256
dark field, 133
intensity, 133
Image, electrical, 397
aperture edge, 700
Image contours, 708
Image defects of electron lenses, *see* Aberrations
Image formation, electron microscope, 87, 672
large aperture, 701
small aperture, 691
Image formation, light microscope, 87
Image multiplier, 50
Image plane, definition, 12, 423
Image point, definition, 11
Image tube, 43, 51, 381
aberrations, 637, 645
Immersion lens, electrostatic, 434, 447
aberrations, 608
Index of refraction, electron, 7, 30
electrostatic field, 352
magnetic field, 354
Induction accelerator, 76, 668
Inelastic scattering, 675
Initial velocities, 41, 126, 207
cathode lens, 577
Insects, 301
Intensity, electron, electron microscope, 262, 704
dependence on specimen, 732
electron microscope gun, 128
Intensity, electron beam, 354
Intermediate image, 88
Inverse feed-back, 234; *see also* Power supply
Ionization gauge, 198

Ion optics, 32, 658
relativistic effect, 658
Ions, 2, 31

Kinescope, 33
Koehler illumination, 86

Laplace equation, **358**, 384, **467**
complex variable, 360
cylindrical coordinates, 375
Fourier solution, 369
separation of variables, 361, 368, 376
Least action, principle, 352, 400
Least time, principle, 351
Light microscope, 83
image formation, 87
Light optics, analogy with electron optics, 4, 345, 350
Limiting aperture, *see* Aperture
Linear accelerator, 72, 660
Living material, observation, 241, 309
Lorentz transformation, 346

Magnetic attraction, electron beam, 594
Magnetic field, 466
bell-shaped, 506
circular loop, 472
determination, 466, 486
in electron lenses, 489
index of refraction, 354
infinite wire, 471
motion of electrons, *see* Trajectories
permanent magnets, 484
scalar potential, 467, 470
solenoid, 473
vector potential, 353, 467
Magnetic lens, aberrations, *see* Aberrations
aberration-free, 605
bell-shaped field, 506
aberration, spherical, 624
aberration, chromatic, 643
cardinal points, 505
coil current regulation, 124, 213
coil design, 144
field determination, 489
field distribution, 139, 152, 490
focal length, 135, 138, 140
high voltage, 650
Magnetic lens, iron-free, 473
long, 15
magnification, *see* Magnification
path determination (approximate method), 513
permanent magnet, 176
refractive power, 16, 136, 508, 612
rotation, 18, 502
rotation-free, 18, 515
short, 16, 514
yoke, magnetic, 175
Magnetic permeability, 466, 480
Magnetic resistance, 143
Magnetic saturation, 480
Magnetic shielding, 481
Magnetomotive force, 135, 478
Magnetron, 528
Magnification, definition, 12, 14
electron lens, 427
magnetic objective, 138, 183
magnetic projector, 149, 183
electron microscopes, 94, 95, 165, **167**, 169, 170, 176, 185
determination in, 258
light optical, of image, 95, 168, 256
maximum useful, 82, 260
uniform accelerating field, 34
Mass spectrograph, 65, 69
Maxwellian distribution, 207
Mechanical model, electrons in magnetic field, 527
Mechanical path plotting mechanism, 417
Membrane, specimen support, 242
Metallurgy, with electron microscope, 331
Metals, electrons in, 2
Microtome, high-speed, 308
Miller indices, 733
Misalignment, 600
Molecules, observation of, 310
Motion of electrons, *see* Trajectories
Multiple scattering, 681, 684
Multiplier, electron, *see* Secondary emission multiplier

Negative electron lenses, 433, 441
Nodal points, 428
Noise, scanning microscope images, 749
Numerical aperture, 82

Object, *see* Specimen
Object plane, 423
Object point, definition, 12
Objective lens, alignment, 252
  tolerances, 216
  depth of field, 155
  magnification, *see* Magnification
Objective lens, electrostatic, 89; *see also* Electrostatic lens
Objective lens, magnetic, 87, 134, 181; *see also* Magnetic lens
  coil design, 144
  field distribution, 135, 139, 154, 490
    measurement, 486
  form factor, 135
  pole pieces, 134, 182, 217
    saturation, 136, 143
Optical distance, 352, 543
Orthicon, 521
Oscillograph, electronic, 33

Paraxial ray equation, 402, 505
  relativity correction, 652
Pencils, imaging, 147, 692, 702
Penetration of electrons, 7, 688
Permanent magnet, 484
Permanent magnet lenses, 176
Photoelectric emission microscope, 116
Photographic emulsions, studies with electron microscope 322
Photographic plate, electron microscope, 259
  voltage characteristics of, 262
Photons, 4
Plastics, 324
Plate chamber, 190
Plural scattering, 681, 685
Point Iconal, 543
Point projector microscope, 106, 738
Pole pieces, *see* corresponding lenses
Porcelain, 331
Potential, electrostatic, 358
  magnetic, scalar, 467, 470
  magnetic vector, 353, 467
Potential distribution, *see also* Laplace equation
  axially symmetric systems, 375, 394
  circular aperture, 383
  coaxial cylinders, 378
  electrolytic plotting, 389
Potential distribution, Liebmann procedure, 386
  slit, 366, 368, 372
  space charge, 396
  two-dimensional systems, 359
    special relations, 371
Power supply, electron microscope, 97, 163, 223
  filament current, 238
  magnetic lens current, 236
  radio frequency, 228
  regulating circuits, 233
  tolerances, 213, 223
  voltage multiplying circuits, 231
Preparation of specimen, 241; *see also* Specimen
Principal planes, 18, 427
  definition, 13
Probe, electron, scanning microscope, 99, 735
  shadow microscope, 110, 736
Projector lens, aberrations, *see* Aberrations
  double, 93, 154
  electrostatic, 90, 93
  magnetic, 88, 147, 181
    coil current, 150
    pole piece, 143, 151
  magnification, *see* Magnification

Quantum mechanics, 345

Radio frequency power supply, 228
Ray equation, 400, 504, 652
  paraxial, 402, 505, 652
Ray paths, *see* Trajectories and Pencils, imaging
Rectifier, *see* Power supply
Reflection of electrons, *see* Trajectories and Electron mirror
Reflection of light, 6
Refraction of electrons, *see* Trajectories
Refraction of light, 6
Refractive power, aperture lens, 445
  electron mirror, 463
  immersion lens, 448
  magnetic lens, 16, 136, 508, 612
  unipotential lens, 438
Regulator, electrical, *see* Power supply
Relativity, special theory, 345
Relativity effect, 30, 650, 667

Relativity effect, electrostatic lens, 651
magnetic lens, 650
Replica technique, surfaces, 247, 330
Resolving power, electron microscopes, 84, **123**, 205, 694, 703, **724**, 735
determination, 727
emission microscope, 741
mass thickness, minimum observable, 158
point emission microscope, 741
point projector microscope, 109, 738
power supply stability, effect, 213, 595
recognition of polygon, 311
relativistic aberration, 215
scanning microscope, 101, 735
shadow microscope, 111, 736
specimen thickness, 156, 720
stray fields, effects, 125, **159**, 217, **597**
Resolving power, eye, 81
Resolving power, lens, 82
Resolving power, light microscope, 83
Resolving power, X-ray shadow microscope, 738
Rest mass, 347
Rotation of image, magnetic lens, 18, 502
Rotational distortion, *see* Aberrations, distortion
Rubber, 313, 325
Rubber model, 418

Sagittal surface, 23, **551**, 579, 637
Saturation, magnetic, objective lens, 136, 143
Scalar potential, magnetic field, 467, 470
Scanning electron microscope, 98
noise considerations, 749
resolving power, 101, 735
surface examination, 102
Scattering of electrons, 7, 674
angular distribution, 681
electron microscope, 87
effect on image, 704
volume scattering, 157, 721
Screen, specimen support, 241
Screen lenses, 433, 441
Secondary emission, 4, 59
Secondary emission multiplier, 59, 369, 520
Secondary emission multiplier, electrostatic, 62, 369
magnetic, 61, 520
orbital beam, 63
reduction of noise, 747
Sectioning technique, 301, 308
Separation of variables, 362, 368, 376
Shadow electron microscope, 106, 736
Shielding, 125, 159, 165, 167, 169, 217, **481**
Short lens, aberration, 616
"Shot noise," scanning microscope, 101, 751
Signal-to-noise ratio, 748
Single scattering, 680, 683
Smokes, 315
Snell's law, 6
Source, electron, *see* Gun
Space charge, 27, 37, 39, 129, **591**, **647**; *see also* Aberrations
lens action, 434
potential distribution, 396
Specimen, electron microscope, *see* class of material
preparation, 241
sectioning, 301, 308
staining, 296, 305
surfaces, reflected electrons, 269
replica, 247, 330
scanning microscope, 102
thickness, 156, 208, 720
Specimen chamber, 185
Specimen stage, 93, 186, 190
Specimen support, 186, 241
Spherical aberration, *see* Aberrations
Spherical aberration coefficient, 604
Spot size, 39, 41
Stability, power supply, *see* Power supply
Staining, electron microscope specimen, 296, 305
Stereoscopic micrographs, 262
Surface examination, *see* Specimen
Symmetry, axial, 375; *see also* Axially symmetric systems
two-dimensional, 359

Tangential surface, 23, **551**, 579, 637
Television, electronic, 33
Thermionic emission, 3
electron microscope study, 113

Thermocouple gauge, 196
Thick lens, 435; *see also* Electron lens
  definition, 13
Thin lens, 435, 514; *see also* Electron lens
  definition, 12
Thomson-Whiddington law, 687
Thread beams, 435, 592
Trajectories, *see also* Deflection of electron beam
  electric field, 8, 400
  electric and magnetic fields, 494
  magnetic field, 14, 493
    axially symmetric, 499
    two-dimensional, 515
  paraxial, 402, 505, 652
  ray equation, 400, 504, 652
Trajectory determination, approximated fields, 412, 513, 535
  differences, method, 407
  Gans method, 412
  graphical methods, 414, 512
  mechanical methods, 417, 527
  rubber model, 418
Transmission, electrons through matter, 7, 674

Uncertainty principle, 349, 695
Uniform fields, electrostatic, 405
  magnetic, 15, 495
Unipotential lens, 437; *see also* Electrostatic lens

Vacuum seals, demountable, 165, 167
Vacuum system, electron microscope 165, 172, **195**
Vector potential, magnetic, 353, **467**, 475
Velocity, electron, *see* Electron
  electron wave, 347
    group, 349
    phase, 347
Velocity spectrograph, 67
Virtual image point, definition, 11
Viruses, 282
  serum reaction, 287
Voltage regulation, electron microscope, electrostatic, 90, 215
    magnetic, 213
Voltage regulator, *see* Power supply
Volume scattering, 157, 721

Wave group, 348
Wave length, electron, 5, 84, 348, **351**
Wave mechanics, 345

X-ray shadow microscope, 111, 738
X-ray shielding, electron microscope, 202, 219